DOCUMENT RESUME

ED 063 047 PS 005 700

AUTHOR	Barr, Rebecca C.
TITLE	Development of a Word Learning Task to Predict Success and Identify Methods by Which Kindergarten Children Learn to Read. Final Report.
INSTITUTION	Chicago Univ., Ill.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-9-E-125
PUB DATE	Feb 71
GRANT	OEG-5-70-0002(010)
NOTE	100p.; [Filmed from best available copy]
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Grade 1; *Kindergarten Children; Learning Processes; Predictive Ability (Testing); *Reading Readiness Tests; *Task Performance; Test Reliability; Test Validity; *Word Study Skills
IDENTIFIERS	Metropolitan Readiness Tests; *Word Learning Tasks

ABSTRACT

The objectives of this study were (1) to construct and evaluate the reliability and validity of word learning tasks for predicting success in learning to read, and (2) to compare the effectiveness of word learning tasks with readiness measures in predicting reading in grade one. Two samples of like-aged boys and girls from urban (76) and suburban areas (105) were selected for the study; one sample was tested with the Word Learning Tasks in May of the kindergarten year and the other in September in first grade. Word learning and reading samples were collected in December and May of first grade for all subjects. Mills' Learning Methods Test scores and readiness information (test scores and teacher ratings) were collected for selected subjects. The combined Tasks were found to be reliable measures, with a moderate degree of validity. A comparison of the Tasks, Words Learned-December, kindergarten teacher ratings, and the Metropolitan Readiness Tests showed that the Metropolitan Readiness Tests were generally the best predictor of reading in May of first grade; the Word Learning Tasks were less effective predictors, and the kindergarten teacher rating was the least effective. Words Learned-December was the most useful measure in indentification of children with low reading skill at the end of the first grade. (Author/DB)

ED 063047

COPY AVAILABLE
BEST

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

OE BR-9-E-125

FINAL REPORT

Project No. 9-E-125

Contract No. OEG-5-70-0002(010)

DEVELOPMENT OF A WORD LEARNING TASK TO PREDICT SUCCESS AND IDENTIFY METHODS BY WHICH KINDERGARTEN CHILDREN LEARN TO READ

Rebecca C. Barr

The University of Chicago

Chicago, Illinois 60637

February 14, 1971

The research reported herein was performed pursuant to a contract with the Office of Education, U. S. Department of Health, Education and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

PS 005700

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

TABLE OF CONTENTS

	Page
LIST OF TABLES	iii
LIST OF FIGURES	v
SUMMARY	vi
Chapter	
I. BACKGROUND AND PURPOSE	1
Problem	1
Purpose	2
II. STANDARDIZATION OF THE WORD LEARNING TASK	3
Methodological Issues	3
Methods	4
Results	11
Summary	35
III. COMPARISON OF THE PREDICTIVE EFFECTIVENESS OF LEARNING AND READINESS MEASURES	37
Purpose	37
Methodological Issues	37
Methods	39
Instruments	40
Subjects	42
Findings	42
IV. SUMMARY	56
Conclusion: Word Learning Tasks	56
Conclusions: Predictive Effectiveness of the Tasks and Other Measures	59
Implications for Practice	60
Future Research	61
APPENDIX	
A. Development of the Word Learning Tasks and Task Materials and Directions	63
B. Mills Learning Methods Test	78
C. Reading and Readiness Instruments	82
D. Number of Subjects Selected and Tested	86
E. Identification of Disabled Readers	88
REFERENCES	93

LIST OF TABLES

Table		Page
1.	Number of Subjects by Sex and Area Assigned to Treatment Groups and Tested in May or September.	9
2.	Means, Standard Deviations, and t-Tests Comparing May and September Samples on Auditory and Visual Task Performance.	14
3.	Means, Standard Deviations, and t-Tests Comparing May and September Samples on Reading Test Performance	15
4.	Discrepancy between Auditory and Visual Task Total Scores Obtained through Group Administration for High, Average, and Low Performing Subjects and Total Sample (N=181)	16
5.	Means, Standard Deviations, and Reliabilities for Form 1 and Form 2 of Auditory, Visual, and Combined Tasks	17
6.	Reliability Coefficients for Word Learning Tests	18
7.	Reliability Coefficients for Auditory, Visual, and Combined Task Learning Retention Scores for Subgroups of Subjects.	19
8.	Means, Standard Deviations, and t-Tests of Difference between Learning-Retention Means of Administration 1 and 2 of Tasks . . .	20
9.	Mean and Standard Deviations for Task Performances for Individual and Group Administration	21
10.	Correlation Coefficients for Group and Individually Administered Tasks (N=76) . . .	22
11.	Correlation Coefficients for Auditory and Visual Mills Scores and Task Scores of 10 and 35 (N=38)	26
12.	Correlation Coefficients for Auditory and Visual Scores for the Mills and Tasks and Reading Skills (N=38).	27

Table		Page
13.	Correlation Coefficients for the Task Learning-Retention Scores and Reading Scores for Three Sample Sizes	28
14.	Correlation Coefficients for Task and Words Learned, December, and Efficiency Scores for May and September Samples	29
15.	Correlation Coefficients for Combined Tasks, December Words Learned, Efficiency and May Reading Skill for May and September Samples	30
16.	Correlation Coefficients for December Words Learned, WRAT May Reading Scores	31
17.	Mean Scores for Tasks and Reading Tests for Subjects Grouped According to Sex, Area, and Task Testing Time	32
18.	Correlation Coefficients of Unstructured Teacher Ratings, Structured Teacher Ratings and Metropolitan, Tasks, December/ WRAT, May/WRAT, Gates-MacGinitie Vocabulary, and Gates-MacGinitie Comprehension (N=70).	41
19.	Number of Subjects and Date of Administration of Tasks, Metropolitan, and Teacher Ratings	43
20.	Correlation Coefficients between Tasks, Words Learned-December, Metropolitan, and Teacher Ratings for Urban and Suburban Samples Tested on the Tasks in May and September	44
21.	Means and Standard Deviations on Tasks, Words Learned-December, Metropolitan, and Teacher Ratings for Urban and Suburban Samples Tested on the Tasks in May and September	45
22.	Correlation Coefficients between May Reading Scores and Tasks, Words Learned-December, Metropolitan, and Teacher Ratings for the May and September Samples . .	46

Table		Page
23.	Correlation Coefficients between May Reading Scores and Tasks, Words Learned-December, Metropolitan, and Teacher Ratings for Boys and Girls	48
24.	Selected Cut-Off Scores on the Task, Metropolitan, Teacher Rating and Words Learned-December for Identification of Adequate Readers (30 or Above on WRAT-May) and Disabled Readers (Below 30 on WRAT-May) (N=108)	52
25.	Combinations of Scores on the Task (Cut-Off Score = 30), Metropolitan (Cut-Off Score = 48), and Teacher Rating (Cut-Off Rating = Medium Low) for the Identification of Adequate Readers (30 or Above on WRAT-May) and Disabled Readers (Below 30 on WRAT-May) (N=108)	54
26.	Percent of Learning and Retention for 24 and 48 Hour Intervals	67
27.	Mean Learning Scores (Total Possible = 25) for Four Groups of Subjects on the Auditory and Visual Tasks	68
28.	Mean Scores for Each Multiple-Choice Check (Possible Score = 5) following Learning on the Auditory and Visual Tasks and for the First and Second Administration	68
29.	Number of Subjects in Each Subgroup (1) Selected for Study, (2) with Completed Learning Task Data, and (3) with Complete First Grade Reading Data	87
30.	Subgroups of Subjects Identified Using Cut-Off Points on the Tasks, Metropolitan, Teacher Rating and Words Learned-December (N=108)	92

LIST OF FIGURES

Figure		Page
1.	Summary Sheet for Scores Recorded from the Test Booklet of Visual Task, Form 1	6
2.	Distribution, Means, and Standard Deviations for Auditory and Visual Task Scores for All Subjects (N=181)	12
3.	Distributions, Means, and Standard Deviations for Auditory and Visual Task Scores for Subgroups of Subjects	13
4.	Distribution, Means and Standard Deviations for Auditory and Visual Scores on the Tasks and the Mills	24
5.	Trends Based on Correlation Coefficients between Readiness and Learning Measures and Reading Assessed during First Grade	50
6.	Distribution of Scores on the Task in Relation to Scores on the WRAT	89
7.	Distribution of Scores on the Metropolitan in Relation to Scores on the WRAT. . . .	90
8.	Distribution of Scores on the Words Learned-December in Relation to Scores on the WRAT	91

SUMMARY

The objectives of the study were 1) to construct and evaluate the reliability and validity of word learning tasks for predicting success in learning to read and 2) to compare the effectiveness of word learning tasks with readiness measures in predicting reading in grade one.

An auditory and a visual word learning task were developed as miniature versions of the learning that a child might later experience. The directions, procedures, and interpretation of responses were clearly defined so that the Tasks might be used by persons with relatively little testing experience. The Tasks were designed so that they could be administered to small groups of children as well as to individuals.

Two samples of like-aged boys and girls from urban (N=76) and suburban areas (N=105) were selected for the study; one sample was tested with the Word Learning Tasks in May of the kindergarten year and the other in September in first grade Word learning and reading samples were collected in December and May of first grade for all subjects. Mills' Learning Methods Test scores and readiness information (test scores and teacher ratings) were collected for selected subjects.

Reliability and Validity

The characteristics measured by the Word Learning Tasks (Tasks) were sufficiently stable over the period of one week to indicate that the combined Tasks are reliable (.929). The Auditory and Visual learning-retention scores evaluated separately are somewhat less reliable (.893 and .847, respectively), but well within the range commonly found for learning tasks. Word learning in group and individual settings is similar.

The validity of the Tasks as assessed by comparison of it with other word learning measures is of moderate degree. The median validity coefficient between the Mills' Learning Methods Test and the Tasks was .541. Similarly, the degree of relation between the Tasks and word learning during the first three months of first grade (Words Learned - December) is moderate: .658 for the May sample, and .609 for the September sample. The Tasks are similar to the Mills' Learning Methods Test and words learned during the first three months of first grade (Words Learned - December) for predicting reading achievement in May.

Predictive Effectiveness

A comparison of the Tasks, Words Learned-December, kindergarten teacher ratings and the Metropolitan Readiness Tests revealed that the Metropolitan Readiness Tests were generally the best predictor of reading in May of first grade. The Word Learning Tasks were consistently (but not significantly) less effective in prediction than the Metropolitan and Words Learned - December; kindergarten teacher rating was generally least effective.

In the identification of children with low reading skil skill at the end of first grade, Words Learned-December, was the most useful measure. No measures obtained prior to reading instruction (Tasks, Metropolitan, teacher rating) were as efficient as Words Learned-December in identifying children who would have most difficulty reading. These findings suggest that the most useful information for identifying children in need of special help can be obtained by the first grade teacher during the first few months of first grade.

I.

BACKGROUND AND PURPOSE

Problem

Many reading readiness tests are presently on the market; these tests measure perceptual and language skills judged to be required in learning to read. Nevertheless, most readiness tests are judged[1] to be insufficiently valid because they are ineffective in predicting reading achievement following one year of instruction. For example, frequently less than 30 per cent of the variation of reading performance at the end of the first grade is accounted for. The failure of readiness tests to predict later reading accurately could have the following source.

Tests that measure component skills may overlook capabilities involved in the process of reading. Most available instruments measure past learning rather than efficiency of acquiring new skills or information. Informal testing in the clinic setting reveals that many children who show the same level of performance in a given readiness skill differ considerably when given the opportunity to learn new skills. Most readiness instruments measure component skills that are expected to combine in some way in reading acquisition. They may be inadequate predictors because they measure separate component skills rather than the integrated process. Miniature versions of learning have face validity because they use the process demanded in later learning. They have been found to be highly predictive of later performance in some areas.[2, 3] Moreover, reading measured at one point in time has been found to be one of the best predictors of

---

[1]R. L. Hillerich, An interpretation of research in reading readiness. Elementary English, 1966, 43, 359-64, 372.

[2]James Wardrop, A learning test approach to predicting classroom performance, ERIC - Research in Education, September, 1963, 3, 28.

[3]K. deHirsch, J. Jansky and W. Langford, Predicting Reading Failure. (New York: Harper and Row, 1966) See Word Recognition II prediction coefficients, pp. 94-97.

reading at a later time.[1] Evaluation is needed to determine if miniature learning tasks are more effective than readiness tests in predicting reading.

In order to compare the predictive effectiveness of learning samples[2] with readiness measures of component skills, it was first necessary to develop and standardize miniature word learning tasks. The first purpose of this report is, therefore, to present information on the standardization of such learning tasks. Standardization entails demonstration that word learning as measured by the learning tasks is a stable characteristic, and that the tasks are reliable. Validation involves comparison of results from the tasks with those obtained from a second learning task and from an estimate of words learned during the first three months of first grade.

Once the learning tasks were standardized, the problem concerning the predictive effectiveness of readiness tests was explored. Word learning, measured by the tasks during the first three months of first grade, a readiness measure, and a teacher rating of readiness were compared on their effectiveness in predicting reading at the end of grade one and in identifying disabled readers.

Purpose

The purpose of the research then was two-fold:

1. To assess the reliability and validity of miniature word learning tasks.

2. To compare learning and readiness on their effectiveness in predicting first grade reading.

The following sections of the report describe studies undertaken to explore problems corresponding to the purposes. In each section, methodological issues relating to the problem will be discussed, samples and procedures will be described, and the findings will be presented. In the final section of the report, I will summarize the findings, draw conclusions, and make recommendations for educational practice and further research.

---

[1] A. I. Gates and J. LaSalle, The relative predictive values of certain intelligence and educational tests together with a study of the effect of educational achievement upon intelligence test scores, Journal of Educational Psychology, 1923, 14, 517-539.

[2] "Miniature versions of learning" and "learning samples" are used interchangeably.

II.

STANDARDIZATION OF THE WORD LEARNING TASKS

Methodological Issues

Reliability.

The reliability of readiness tests is frequently assessed using a split-half or odd-even item comparison. This procedure gives an estimate of test consistency. It is not, however, sensitive to the child's performance from time to time. Because the performance of young children is greatly influenced by attentional factors, a test may be reliable, but the characteristic measured may not be stable over even a short period of time.

Thus, in the standardization of the miniature word learning tasks (hereafter called Tasks), the test reliability is assessed using an index of the stability of the characteristic. Two forms of the Tasks were constructed to measure word learning with a time period of a week intervening. One week was selected since it is short enough so that no real change in learning would be expected, but sufficiently long to test the stability of the characteristic over a time interval. If the stability of the characteristic measured is shown to be high, then it can be assumed that the test is reliable. If, however, the characteristic is not stable over time, then instability could reflect the nature of the characteristic or the unreliability of the test.

Validity

Although some learning tests are commercially available, they tend to be constructed as associative learning tasks rather than as miniature versions of the reading instruction that the child would later receive.[1] Only one test involves different instructional methods that resemble instruction: the Mills Learning Methods Test.[2] This test was constructed to be used with children who had received some instruction and were experiencing difficulty reading, not with pre-readers.

---

[1]Helen A. Murphy and Donald D. Durrell, Murphy-Durrell Reading Readiness Analysis. (New York: Harcourt, Brace and World, Inc., 1965).

[2]Robert E. Mills, Learning Methods Test (Ft. Lauderdale, Fla: Mills Center, Inc., 1955).

Several problems made me think this test would be inappropriate for pre-readers. The sample of ten words to be learned each instructional session seemed overwhelming to some pre-readers tested in pilot research. The directions are loosely specified and tend to vary from teacher to teacher; therefore, replication of the learning situations would be difficult. Forty words unknown to the child are randomly assigned to each of four methods: auditory, visual, kinesthetic, and combination. Frequently, words assigned to the auditory method are phonetically irregular. Such word-selection procedures make group teaching and testing impossible. More important, random assignment of words to methods is an unnecessary constraint for pre-readers.

For these reasons Mills' test was not used as the learning task. Nevertheless, since it is the only existing learning test which reflects first grade instructional methods, the learning measured by it was compared with that measured by the Tasks constructed for pre-readers.

In addition, I selected a second criterion to examine the validity of the Tasks: the words that each child learned during the first three months in first grade. The child was tested on this word sample from all words taught to him and two scores were obtained: number of words learned (hereafter, referred to as Words Learned, December), and efficiency of learning (percentage of words learned of those that had been presented to him).

Methods

The Tasks were designed to measure learning a set of words using two methods, auditory and visual. The Tasks are miniature versions of the learning that a child might encounter during his first year of instruction. Two forms were constructed for each method. Directions were designed so that the Tasks could be administered to children individually or in small groups.

In order to assess the stability of learning on the Tasks, two sets of scores were obtained for each child using one of the following procedures:

(1) Form 1 - Form 2.
Form 1 (Auditory and Visual) and Form 2 (Auditory and Visual) were administered a week apart to groups of six children,

(2) Group-Individual.
One form of the tasks (Auditory and Visual) was administered to a child individually and the second form, a week

later (earlier) to the child in a group with five other children.

Given this data, the stability of learning over a week period was assessed for group versus group learning and for group versus individual learning.

Validity of the tasks was determined by the following procedures:

(1) Tasks-Mills.
Learning on the Tasks (Auditory and Visual) was compared with learning on the Mills (Auditory and Visual)

(2) Tasks-Words Learned, December.
Learning on the Tasks was compared with the words learned during the first three months of first grade.

In addition, the learning measures were compared on their effectiveness to predict reading at the end of first grade.

Instruments

Word Learning Tasks. Through a series of pilot studies, the Word Learning Tasks were developed. The Tasks were designed to measure learning words using two methods, Auditory and Visual, that are miniature versions of the learning that a child might encounter during his first year of instruction.

The Auditory method incorporates procedures common to a "phonics" or "linguistics" reading program. Pilot research showed five words to be an appropriate sample for pre-readers. Five words were selected that entailed the learning of six sounds. For example, at, hat, sat, mat, and bat, were used on one form. The learning sequence includes the following three activities: discriminating letter sounds; learning letter-sound associations; and blending sounds using the initial consonant substitution method.

The Visual method is similar to the sight-word learning approach used during the first year of some basal reader programs. Five words, systematically varied in length and configuration, are taught using the following three activities: associating word form with prior experiences with the word; discriminating and matching of word forms visually; and identifying the unique features of word forms.

PS 005700

The Auditory and Visual Tasks are designed to be as comparable as possible. Total teaching time for each method is 15 minutes. Teaching is followed by four review-test trials. Testing is done using (1) a multiple-choice word-selection format (using the other words in the sample as foils) four times following the teaching and one time during a retention check twenty-four hours later, and (2) an individually administered production test of the word name in response to the printed symbol, once following the four multiple-choice trials after learning, and once twenty-four hours after the learning. Figure 1 shows the summary sheet for the test booklet of Visual Task, Form 1. A total score of 20 is possible on the multiple-choice items following learning (five words tested during each of the four trials), and a score of five during the production test, yielding a total possible learning score of 25. A total score of five from the production test and five from the multiple-choice test is possible from the twenty-four hour later retention check, making a total possible retention score of 10.

FIGURE I

SUMMARY SHEET FOR SCORES RECORDED FROM THE TEST BOOKLET OF VISUAL TASK, FORM 1

Word Learning Task
Visual Method, Form 1

	Learning Score					Retention Score	
	Multiple-Choice				Naming	Naming	Multiple-Choice
bus	1	1	1	1	1	1	1
tree	1	1	1	1	1	1	1
girl	1	1	1	1	1	1	1
house	1	1	1	1	1	1	1
children	1	1	1	1	1	1	1
Subtotal Scores				20	5	5	5
Total Learning				25		Total Retention	10
Total Learning-Retention							35

The directions given by the teacher-researcher are standard. For further details on pilot research used in the development of the Tasks, and for examples of the Task directions and materials, see Appendix A. Two forms were constructed for the Auditory and the Visual Tasks. The procedures were designed so that they can be administered either individually or to a small group of children.

Mills Learning Methods Tests. The Learning Methods Tests[1] developed by Robert Mills entails four instructional methods: Auditory, Visual, Kinesthetic, and Combined. Two methods, Auditory and Visual, were selected to be compared with the Auditory and Visual Tasks. Twenty words selected from the pre-primer level of the Mills word cards were assigned to the Auditory and Visual Methods for each child tested. (Directions specified by Mills are shown in Appendix B.) Fifteen minutes was allowed for teaching the ten words. Learning was checked by having the child name the ten words. Retention was checked twenty-four hours later using the same procedure.

Words Learned, December. Words learned by children during the first three months of first grade were measured in December, 1969. Teachers were contacted in late November to determine the page where each child was reading in the reading series. A sample of words was systematically selected (every n^{th} word) from the words to which the child had been exposed in his reader. Additional words taught by the teacher (e.g., colors, numbers) were also included. The sample varied from class to class and from child to child. The number of these words that the child was able to pronounce was used to estimate the number of words that he had successfully learned and learning efficiency. For example, if a child had been exposed to 50 words and learned 25, the number of words learned would be 25 and his efficiency 50 per cent. If he learned 25 out of 25 words, the number of words learned would be the same, but his efficiency would be 100 per cent.

Reading Tests. Reading was measured in December and May of first grade using the Wide Range Achievement Test, Reading Section (WRAT)[2] and in May of first grade using the Gates-MacGinitie Reading Test, Primary 1, Form A, Vocabulary

---

[1]Mills, loc. cit.

[2]J. F. Jastak and S. R. Jastak, Wide Range Achievement Tests. (Wilmington, Del.: Guidance Associates, 1965.)

and Comprehension tests,[1] an informal phonics test, and a test of words learned by May (obtained by testing procedures similar to those for words learned by December.) The Gates-MacGinitie tests and the WRAT are similar in validity and reliability (see Appendix C); however the scores from the Gates-MacGinitie tests are more influenced by guessing and picture cues than are those from the WRAT. The WRAT, a pronounciation test of words in isolation, seems to be a more valid estimate of word recognition, whereas the Gates-MacGinitie tests appear to tap a wider variety of reading-related skills, such as use of picture cues and sentence context in addition to word recognition. The phonics test consists of 15 nonsense syllables to be pronounced. A score was given for each letter-sound association given and vowel marker observed, for a total possible score of 45. A sample of the answer sheet for recording responses to the WRAT, Words Learned, and the Phonics Test is shown in Appendix C.

Subjects

Two samples of like-aged boys and girls from urban and suburban areas were selected; one sample was tested in May of the kindergarten year and the other, in September in first grade. Urban subjects tested in May during kindergarten came from two public elementary schools in a lower class to middle class, largely Negro innner-city neighborhood. The suburban group tested in May during kindergarten came from two public elementary schools in a middle- to upper-middle class Caucasian suburban neighborhood.

Similarly, two different samples of children were tested with the Task in September in first grade. The urban group came from two schools in the same inner-city school district as the May sample. One of the schools was in a predominantly lower to lower-middle class, largely Negro neighborhood; the other in a lower to lower-middle class neighborhood where a majority of the children came from Spanish-speaking homes. The suburban sample tested in September came from two public elementary schools in lower-middle to upper-middle class Caucasian neighborhoods. The latter group came from a different suburban area than the suburban sample tested in May.

Children given the Tasks during May of their kindergarten year were assigned to two test comparison groups and those given the Tasks during September in first grade were assigned to the three test comparison groups. Table 1 shows the five samples, lists the learning tests given to each, and gives the number of subjects by sex and area for

---

[1]Arthur I. Gates and Walter H. MacGinitie, Gates-MacGinitie Reading Tests (New York: Teachers College Press, Teachers College, Columbia University, 1965).

whom both learning test data and follow-up reading data was complete.

TABLE 1

NUMBER OF SUBJECTS BY SEX AND AREA ASSIGNED TO TREATMENT GROUPS AND TESTED IN MAY OR SEPTEMBER

Date Tested		Test Comparison Groups: Form 1 and Form 2		Group and Individual		Tasks - Mills		
		Urban	Suburban	Urban	Suburban	Urban	Suburban	Total
May 1969	Boys	10	10	8	10	0	0	
	Girls	4	10	11	8	0	0	
	Total	14	20	19	18	0	0	71
Sept. 1969	Boys	6	10	7	12	8	12	
	Girls	7	10	8	12	7	11	
	Total	13	20	15	24	15	23	110
Group Totals		67		76		38		181

A total of 96 children was tested in May during their kindergarten year, and randomly assigned to either the group where Form 1 was compared with Form 2 (Form 1 - Form 2) or the administration of the Tasks in a group-setting was compared with that in an individual setting (Group-Individual). Learning Task data were incomplete for 13 subjects because of frequent absence, and first grade reading scores were not obtained for 12 subjects who had transferred to other schools. This left a total of 71 children tested in May with complete learning test and follow-up reading data.

One hundred twenty-six children were selected for testing in September during the first grade and assigned to the Form 1 - Form 2 sample, the Group-Individual sample, or the sample in which learning on the Tasks was compared with that on the Mills (Tasks - Mills). Learning test data were incomplete for ten subjects who were frequently absent. First grade reading scores were unavailable for six children who were absent or who had moved from the area. Complete data were obtained for 110 subjects tested

in September of first grade.

For the comparison of Form 1 of the Tasks with Form 2, complete data were obtained for 67 subjects (34 tested in May of Kindergarten, and 33 tested in September of first grade). Scores were complete for 76 subjects given the tasks individually and in a group (37 tested in May, 39 in September). The Tasks and the Mills Learning Methods Tests were administered to 38 subjects in September of first grade. (Detailed information concerning subject drop-outs is contained in Appendix D.)

Procedures

Learning test information was collected during the last two weeks of May of kindergarten for the May groups and during the last two weeks of September of first grade for the September groups. All children participated in four learning sessions and four retention checks (either both forms of the Auditory and Visual Tasks or one form of the Auditory and Visual Tasks and the Auditory and Visual tests of the Mills). If a child was absent on the first day of testing, or if he already knew some words on the task, he was replaced by a previously specified alternate child of the same sex. The retention checks occurred twenty-four hours after the learning session. The series of four learning tasks was usually administered within a two-week period. The following represents a typical schedule:

First Week:	Monday:	Administer Form 1 of Visual Task
	Tuesday:	Check retention of Form 1, Visual Task; Administer Form 1, Auditory Task
	Wednesday:	Check retention of Form 1, Auditory Task
Second Week:	Monday:	Administer Form 2, Visual Task (or Mills)
	Tuesday:	Check retention of Form 2, Visual Task (or Mills); Administer Form 2, Auditory Task (or Mills)
	Wednesday:	Check retention of Form 2, Auditory Task (or Mills)

Each learning session required approximately 40 minutes for teaching and testing. Retention checks took approximately 10 minutes. Total testing time for each child involved approximately 2 hours and 40 minutes.

Forms, methods, and tests were counterbalanced in administration to control for sequence effects. Research assistants doing the teaching were clinicians with teaching experience who had been specially trained and observed

during task administration to insure that directions were appropriately followed.

The data were summarized statistically using mean scores and standard deviations, correlation coefficients (product-moment formula) and reliability coefficients; t-test comparisons were used to determine the statistical significance of differences between mean scores.

Results

All children in the three test comparison groups received at least one group administration of the Tasks. Before reliability and validity data are presented, descriptive statistics and score distributions for the total group and subgroups will be shown and discussed in order to determine if the tasks are too easy or difficult for selected groups of subjects and to evaluate the degree and nature of the relationship between the Auditory and Visual Task scores.

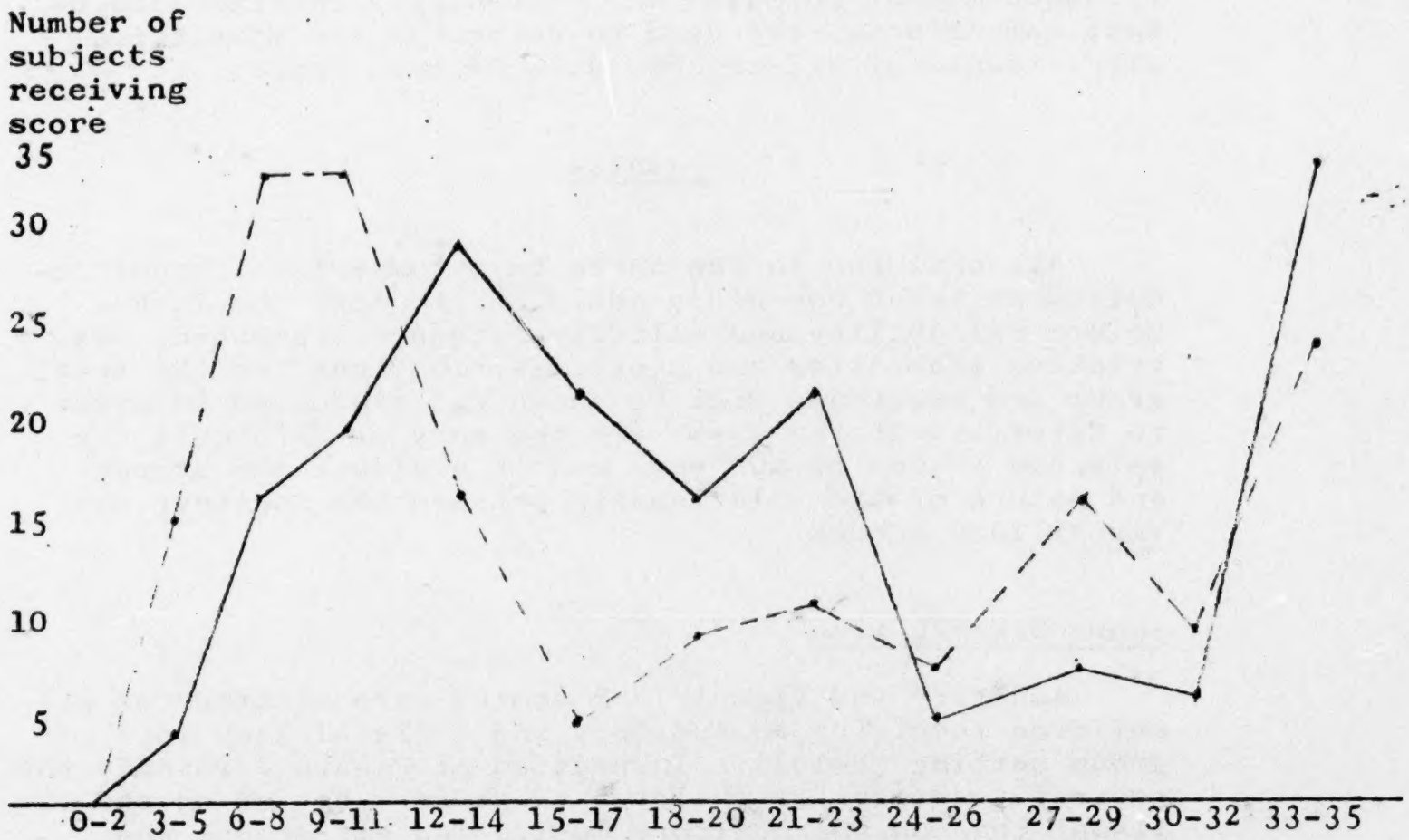
Score Distribution

Auditory and Visual Task scores were plotted for all children receiving an Auditory and a Visual Task in a group setting (N=181). Inspection of Figure 2 reveals the general tendency for children to perform better on the Visual than on the Auditory Task. The difference was found to be statistically significant ($t=7.89$, $p < .001$). Since 25 of the possible 35 scores are obtained using five-choice items, a score of five represents chance expectancy. For both Tasks the distribution of scores is bimodal, indicating that children tend to be either good or poor learners on these tasks. On the Visual Task, there is a tendency for some scores to cluster near the highest possible score; the Visual Task apparently is not difficult enough for some pre-readers.

The distributions of Auditory ar Visual Task scores for subgroups of subjects are shown n Figure 3. The ceiling effect on the Visual task is more apparent for suburban girls tested in September. Comparison of Auditory and Visual Task score distributions and mean scores supports the common finding that girls achieve significantly better than boys in reading ($t = 6.87$ $p = < .001$; $t = 6.76$, $p = <.001$,

FIGURE 2

DISTRIBUTIONS, MEANS, AND STANDARD DEVIATIONS
FOR AUDITORY AND VISUAL TASK SCORES FOR ALL SUBJECTS (N=181)



Scores on Tasks:	Auditory (---)	Visual (——)
Means:	17.23	19.75
Standard Deviations:	10.42	9.48
Correlation:	.831	

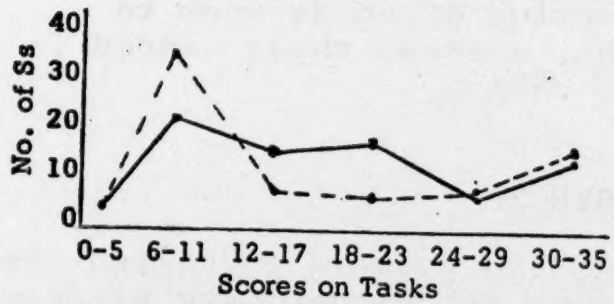
respectively) In keeping with many reported findings, the Auditory and Visual learning of suburban children is signinicantly better than that of urban children ($t = 6.10$, $p<.001$; $t = 4.83$, $p<.001$, respectively.)

The sample of children tested in September of first grade performed significantly better than children who were an average of four months younger when tested in kindergarten. The difference is attributable to improvement on the Visual Task rather than on the Auditory Task. The mean scores for the May and September administration of the

FIGURE 3

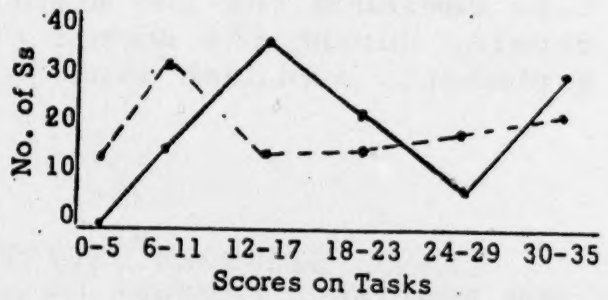
DISTRIBUTIONS, MEANS, AND STANDARD DEVIATIONS
FOR AUDITORY AND VISUAL TASK SCORES
FOR SUBGROUPS OF SUBJECTS

May Sample (N=71)



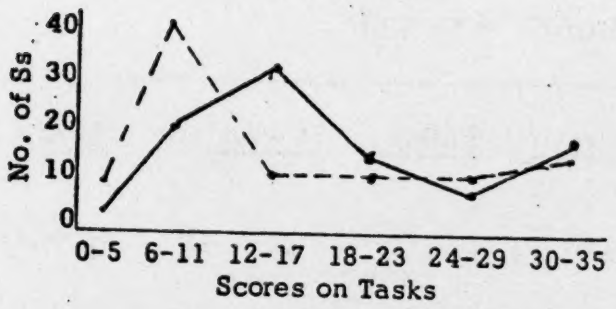
	Mean	S.D.	
Auditory	16.61	10.32	Correlation
Visual	18.03	9.68	.830

September Sample (N=110)



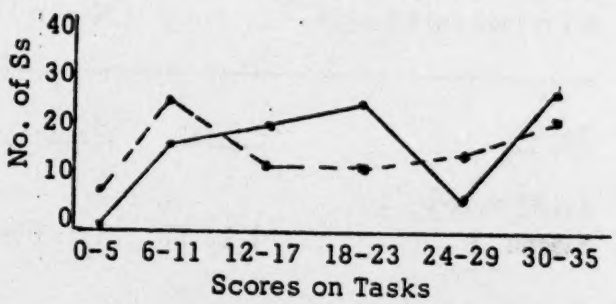
	Mean	S.D.	
Auditory	17.63	10.52	Correlation
Visual	20.86	9.22	.837

Boys (N=93)



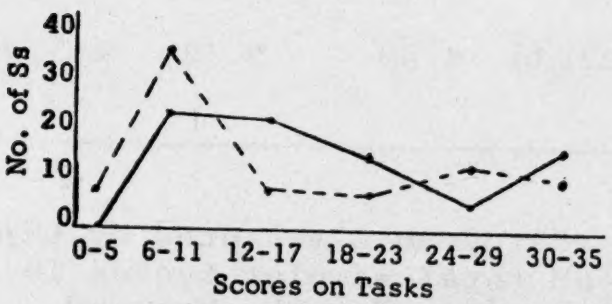
	Mean	S.D.	
Auditory	15.62	10.13	Correlation
Visual	18.28	9.28	.832

Girls (N=88)



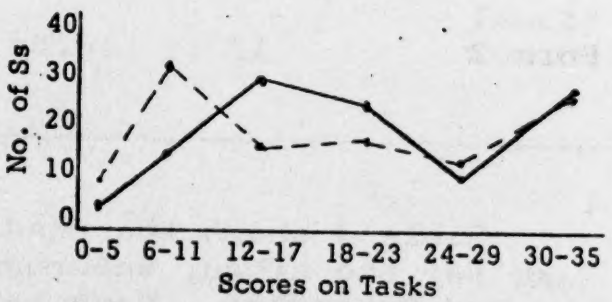
	Mean	S.D.	
Auditory	18.92	10.52	Correlation
Visual	21.31	9.49	.821

Urban (N=76)



	Mean	S.D.	
Auditory	15.53	9.78	Correlation
Visual	18.46	9.34	.808

Suburban (N=105)



	Mean	S.D.	
Auditory	18.46	10.74	Correlation
Visual	20.69	9.51	.842

Auditory (- - -)

Visual (——)

Visual Task are significantly different ($t = 6.03$, $p = <.001$), whereas the difference for May and September administration of the Auditory task just barely achieves statistical significance ($t = 2.08$, $p = <.05$). This finding was replicated in a comparison of 34 children given both Form 1 and 2 in May with 33 children similarly tested in September. Mean scores on the Auditory and Visual Tasks are shown for these May and September subjects in Table 2. The capabilities demanded for the sight learning of words seem to develop during the summer months, whereas those needed in a phonetic approach show little change.

TABLE 2

MEANS, STANDARD DEVIATIONS, AND t-TESTS COMPARING MAY AND SEPTEMBER SAMPLES ON AUDITORY AND VISUAL TASK PERFORMANCE

Time of Task Administration:	May (N=34)		Sept. (N=33)			
Task	Mean	S.D.	Mean	S.D.	t-Value	Signif.
Auditory - Adm. 1	15.32	10.68	16.24	11.03	.52	$p > .05$
Auditory - Adm. 2	17.06	11.22	17.91	11.60	1.03	$p > .05$
Visual - Adm. 1	17.56	10.44	20.88	9.77	4.28	$p < .01$
Visual - Adm. 2	17.41	10.06	22.36	8.99	6.54	$p < .01$

Table 3 shows the reading achievement measured by three tests for the urban, suburban, and total samples tested in May and September. Mean reading scores for the May and September samples do not differ on two of the three reading measures. It is judged that differences on the Auditory and Visual Tasks are not attributable to sample differences. Rather, maturity and/or experiences occuring during the summer appears to increase ability for sight-word learning but not for phonetic learning of words.

TABLE 3

MEANS, STANDARD DEVIATIONS, AND t-TESTS COMPARING MAY AND SEPTEMBER SAMPLES ON READING TEST PERFORMANCE

Time of Task Administration:	May (N=34)		Sept. (N=33)			
Reading Test	Mean	S.D.	Mean	S.D.	t-Value	Signif.
WRAT	39.21	9.68	39.79	8.17-	.80	$p > .05$
Gates - Voc.	33.94	12.46	33.64	11.58	.35	$p > .05$
Gates - Comp.	19.91	9.28	18.30	8.96	2.18	$p \leq .05$

Comparison of Auditory and Visual Task Scores

The reader should note in Figure 3 the magnitude of correlation coefficients found for the subgroups in auditory and visual learning. Learning ability seems to be a general factor rather than several factors related to modality of perception as indicated by the fact that children who performed well on the Visual Task also performed well on the Auditory Task. Similarly, those who achieved poorly on one Task, achieved poorly on the other.

Some children did, however, perform differently on the Auditory than on the Visual Task. Discrepancy scores derived from the difference between scores on the Auditory and Visual Tasks are shown in Table 4. Most children show little discrepancy, and the direction of the discrepancies is towards a higher visual score.

Discrepancies occur about equally for subjects with high, average or low combined scores. Because of the ceiling effect of the tests, more children with high Combined Task scores show zero discrepancy (9 of the 14 children). Generally a discrepancy between modalities of learning is as uncommon for good readers as for poor although there seems to be a slight tendency for average and low subjects to perform higher on the Visual Task.

TABLE 4

DISCREPANCY BETWEEN AUDITORY AND VISUAL TASK TOTAL SCORES OBTAINED THROUGH GROUP ADMINISTRATION FOR HIGH, AVERAGE, AND LOW PERFORMING SUBJECTS AND TOTAL SAMPLE (N=181)

	Discrepancy between Task Scores									
	Higher Auditory				No Diff.	Higher Visual				N
	16-20	11-15	6-10	1-5	0	1-5	6-10	11-15	16-20	Total
Total Sample	1	2	8	37	20	65	30	15	3	181
Low Subjects Combined Score=0-24	0	0	2	18	4	26	12	6	0	68
Avg. Subjects Combined Score=25-48	1	0	2	12	2	18	10	8	3	56
High Subjects Combined Score=49-64	0	2	4	7	14	21	8	1	0	57

Modality strengths as measured by perceptual tests in a clinical setting are frequently found to be unstable on the basis of a second testing. To see whether the modality strengths indicated by the discrepancy scores were stable over a one-week period, discrepancy scores were compared for all subjects receiving Form 1 - Form 2 administration of the Tasks (N=67). The resulting coefficent of correlation (r=.133) shows that preferred mode of learning identified through the first administration of Auditory and Visual Tasks is not related to that from a second administration. This finding questions the stability of modality preference as measured by the Tasks.

Although children are somewhat more successful on the Visual than on the Auditory Task, performance on the two Tasks is highly correlated. Evidence suggests that a general learning ability is tapped rather than two distinct perceptual modalities. Modality preference as estimated by discrepancy scores was not found to be stable across two Task administrations.

Reliability of the Word Learning Tasks

Form 1 vs. Form 2 Comparisons. The stability of performance on the Tasks was evaluated over an interval of approximately one week using comparable forms of the Tasks. Sixty-seven subjects were tested in groups on Form 1 and Form 2. Forms were counterbalanced to correct for sequence effects. A period of approximately one-week intervened between administrations. Means and standard deviations for Forms 1 and 2 and estimates of stability are shown in Table 5. Mean scores for learning, for retention, and for learning and retention were compared for Forms 1 and 2 of the Auditory, Visual and Combined Tasks using t-tests to estimate the significance of the difference. The means of the visual retention scores were significantly different at the .001 level. The difference was also reflected in the Visual learning-retention means ($p < .05$) and the combined learning-retention means ($p < .05$). No other means were found to be significantly different at the .05 level of confidence or less. This shows that the scores from Form 1 are generally comparable to those from Form 2 of the Auditory and Visual Tasks with the exception of the Visual retention test.

TABLE 5

MEANS, STANDARD DEVIATIONS, AND RELIABILITIES FOR FORM 1 AND FORM 2 OF AUDITORY, VISUAL, AND COMBINED TASKS (N=67)

Task	Form 1 Mean	Form 1 S.D.	Form 2 Mean	Form 2 S.D.	Stability (Reliability)
Auditory					
Learning (25)	12.88	7.69	12.96	8.11	.875
Retention (10)	3.85	3.50	3.82	3.55	.756
Learning-Retention (35)	16.73	10.87	16.78	11.28	.893
Visual					
Learning (25)	14.46	7.40	14.19	6.94	.827
Retention (10)	4.81	3.31	5.58	3.04	.695
Learning-Retention (35)	19.27	10.30	19.77	9.67	.847
Combined					
Learning (50)	27.34	14.50	27.15	14.40	.912
Learning-Retention (70)	36.00	20.38	36.55	20.17	.929

The retention scores were least stable for both the Visual and the Auditory Tasks, whereas the respective learning retention scores were most stable. The Auditory learning-retention score shows an acceptable level of stability. The Visual learning-retention score approaches an acceptable level of stability. The Combined (Auditory and Visual) learning-retention score shows adequate stability (r=.929). Since the learning-retention scores on the Visual, Auditory and Combined tasks are found to be highly stable, the tasks are judged to possess at least a comparable degree of reliability.[1]

Reliability of learning tasks are frequently not included in reports of studies using them. However, for those test manuals and research articles reporting such information, reliability coefficients are shown in Table 6. Comparison of the levels of reliability shown in Tables 5 and 6 reveals that the Visual and Auditory Tasks are somewhat less reliable than the Mills Visual and Auditory Tests. The Mills sample was inadequately described but appeared to be selected from a range of grade levels, and, possibly, reading levels. Such a selection procedure would serve to increase the magnitude of test-retest coefficients.

TABLE 6

RELIABILITY COEFFICIENTS FOR WORD LEARNING TESTS

Mills Learning Methods Test	Reliability Coefficients	
Methods:	N = 30[2]	N = 38[3]
Visual	.969	.850
Auditory	.970	.802
Kinesthetic	.908	
Combination	.954	
Murphy-Durrell Reading Readiness Analysis[4]		N = 200
Learning Rate Subtest		.88

[1]Reliability and stability will be used interchangeably, and refer to a characteristic of a test.

[2]Mills, loc. cit. (Test-retest correlation coefficients)

[3]Odd-even split-half correlation coefficients corrected by Spearman-Brown Prophesy Formula. (Sample from the present study).

[4]Murphy and Durrell, loc. cit.

The Auditory, Visual, and Combined Tasks were found to be the most reliable when the learning-retention score was used. Therefore, this score will be used in the further analyses reported.

Reliability (stability) coefficients for subgroups of subjects are shown in Table 7. Generally, coefficients are higher for girls than boys, for children tested in September than in May and for urban than suburban children. The finding from the last comparison was unexpected. Given the higher noise levels and number of visual distractions during the learning in urban schools as opposed to suburban schools, it was anticipated that the urban children would show less stable learning. Closer inspection of the reliabilities shows that the Auditory and Visual Tasks differ for urban and suburban children: learning on the Visual Task was more stable for urban children whereas learning on the Auditory Task was more stable for suburban children. Nevertheless, the differences between correlation are not statistically significant ($p > .05$).

TABLE 7

RELIABILITY COEFFICIENTS FOR AUDITORY, VISUAL, AND COMBINED TASK LEARNING RETENTION SCORES FOR SUBGROUPS OF SUBJECTS

Subgroup	N	Auditory	Visual	Combined
May	34	.864	.857	.909
September	33	.921	.828	.951
Boys	36	.863	.849	.918
Girls	31	.925	.829	.935
Urban	27	.841	.904	.952
Suburban	40	.927	.817	.920

In order to determine whether the first administration of the Tasks facilitated learning on the second administration of the Tasks, mean scores from the two administrations were compared. As shown in Table 8, there is a mean increase of 1.45 scores on the Auditory Task and .66 scores on the Visual task; the improvement was found to be statistically significant ($p < .001$ and $p < .01$, respectively). Some learning to learn was demonstrated especially on the Auditory Task. This confirms the tentative findings from the pilot research presented in Appendix A.

TABLE 8

MEANS, STANDARD DEVIATIONS, AND t-TESTS OF DIFFERENCE BETWEEN LEARNING-RETENTION MEANS OF ADMINISTRATION 1 AND 2 OF TASKS (N=67)

	Administration 1		Administration 2			
Task	Mean	S.D.	Mean	S.D.	t-Value	Signif.
Auditory (35)	16.03	10.77	17.48	11.33	8.02	$p < .001$
Visual (35)	19.19	10.17	19.85	9.80	3.09	$p < .01$
Combined (70)	35.22	20.03	37.33	20.39	11.27	$p < .001$

Group vs. Individual Administration.

Children examined in clinics or by school psychologists are frequently tested individually. Generalization is commonly made from learning that occurs in an individual setting to group learning in the classroom. In order to estimate the stability of learning across group and individual settings, 76 children were given one administration of the learning tasks individually and another in a group with (usually) five other children. Test conditions were counterbalanced, half of the children receiving the individual administration first and half the group receiving it last. Table 9 shows the means and standard deviations of children tested individually and in a group. The Auditory and Visual average learning scores are similar under the two conditions as are the Visual retention and learning-retention scores. The group condition, however, seemed to have a slight facilitating effect on Auditory retention. Reliabilities computed for the Auditory and for the Visual Tasks adminis-

tered under the two conditions are .858 and .790 respectively. These coefficients are not significantly lower ($p > .05$) than those reliabilities obtained for Auditory and for Visual Tasks when both were administered under group conditions (.893 and .847, respectively). Thus, the tasks appear to be reliable across group and individual administrative conditions.

TABLE 9

MEAN AND STANDARD DEVIATIONS FOR TASK PERFORMANCES FOR INDIVIDUAL AND GROUP ADMINISTRATION (N=76)

	Group		Individual			
	Mean	S.D.	Mean	S.D.	t-Value	Signif.
Auditory WLT						
Learning (25)	13.92	7.28	13.76	7.84	.84	$p > .05$
Retention (10)	4.66	3.38	4.33	3.19	2.35	$p < .05$
Learning-Retention (35)	18.58	10.35	18.09	10.59	2.48	$p < .05$
Visual WLT						
Learning (25)	14.79	6.63	14.95	6.75	.75	$p > .05$
Retention (10)	5.66	3.26	5.87	3.31	1.30	$p > .05$
Learning-Retention (35)	20.45	9.57	20.82	9.62	1.39	$p > .05$

An interesting question was raised by the following findings: the correlation between the Auditory and Visual tests administered in the group setting was .856 (Table 10). In the individual setting the correlation was significantly lower, .718 ($t = 5.31$, $p < .01$). Modality discrepancies were greater for the learning in an individual setting than for the group administration. Although the methods, procedures, and total time used in the individual and group learning were similar, such

as
factors /reinforcement from the examiner, the child's responding, or pacing geared to the child's reactions may have allowed one modality (auditory or visual) to be used more successfully in an individual than in a group setting. Further evidence concerning the stability of such discrepancies in the individual setting is needed before any conclusions can be drawn concerning effect of individual learning. Generally, however, the findings indicate that on the Auditory and Visual Tasks mean achievement in group and individual settings is similar.

TABLE 10

CORRELATION COEFFICIENTS FOR GROUP AND INDIVIDUALLY ADMINISTERED TASKS (N = 76)

	Group		Individual	
	Auditory	Visual	Auditory	Visual
Group - Auditory	--			
Group - Visual	.856	--		
Indiv. - Auditory	.853	.789	--	
Indiv. - Visual	.774	.790	.718	--

The major findings concerning the reliability of the Tasks are as follows. The characteristic measured by the Combined (Auditory and Visual) Tasks is sufficiently stable over the period of one week to indicate that the Combined Task is reliable. The Auditory and Visual learning-retention scores are somewhat less reliable, but within the acceptable range. Learning in group and individual settings was similar; however, discrepancy between Auditory and Visual learning was somewhat higher in the individual setting. The Tasks appear to be measuring a stable characteristic.

Concurrent and Predictive Validity

Several measures of word learning were compared with the Tasks to determine if similar abilities were being measured and to assess relative predictive validity. The Mills Learning Methods Test, Auditory and Visual subtests, and the Tasks were administered to a group of children. In a second analysis, an estimate of words learned during the first three months in first grade (Words Learned, December) was

obtained for all children (N = 181). This word learning estimate was compared with learning measured by the Combined Task.

Tasks vs. Mills Comparison. Thirty-eight subjects from urban and suburban areas were tested on the Mills Auditory and Visual tests and on the Tasks in September of the first grade prior to reading instruction. On each of the Mills tests, a total score of 20 was possible, 10 if all words were correctly named at the end of the learning period and 10 if all words were correctly named a day later. On the Task a total score of 35 was possible. During the learning session, a score of 20 was possible if the answers to all multiple-choice items were correctly selected and five if the five words presented were correctly named, making the total possible learning score 25. On the retention test a day later, a score of five was possible if the five words presented were correctly named by the child and five if the child could select from the five words the word named by the examiner, making the total possible retention score 10. The two scores on the Tasks that are most comparable to those on the Mills are those where the words are named by the child (five from learning and five from retention) yielding a possible total score of 10. In the first comparison of the Tasks with the Mills this score of 10 is used. In the predictive analyses the more reliable score of 35 is used.

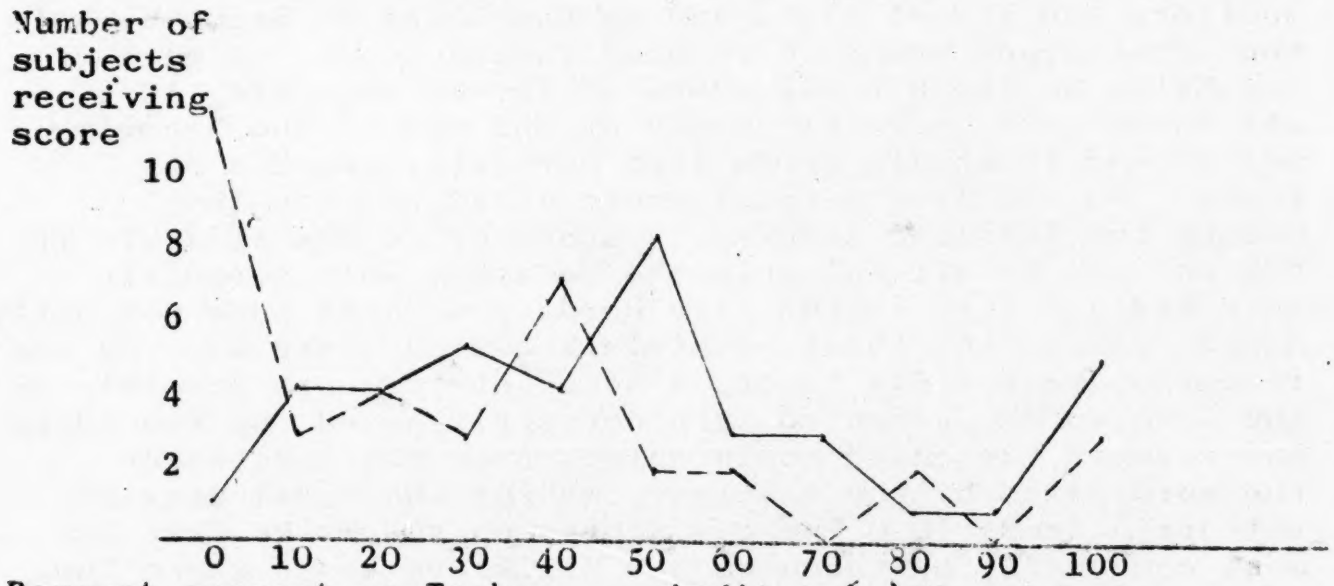
The distributions of scores from the Mills (possible score = 20) and from the Tasks (possible naming score = 10) are shown in Figure 4. When the reader compares these scores he should keep in mind that 100 per cent on the Mills means that 10 words have been learned whereas 100 per cent on the Tasks means that five words have been learned. Roughly one-sixth of the subjects can cope with six or more words on the Mills. However, fewer children than this learn five words on the Tasks.

Three children on the Auditory Task and five on the Visual Task received perfect scores. Although the Tasks seemed to possess sufficient range for most children, there are some prereaders for whom it is too easy. As demonstrated by the Mills, most prereaders (roughly 80 per cent) however, can only cope with five or fewer words.

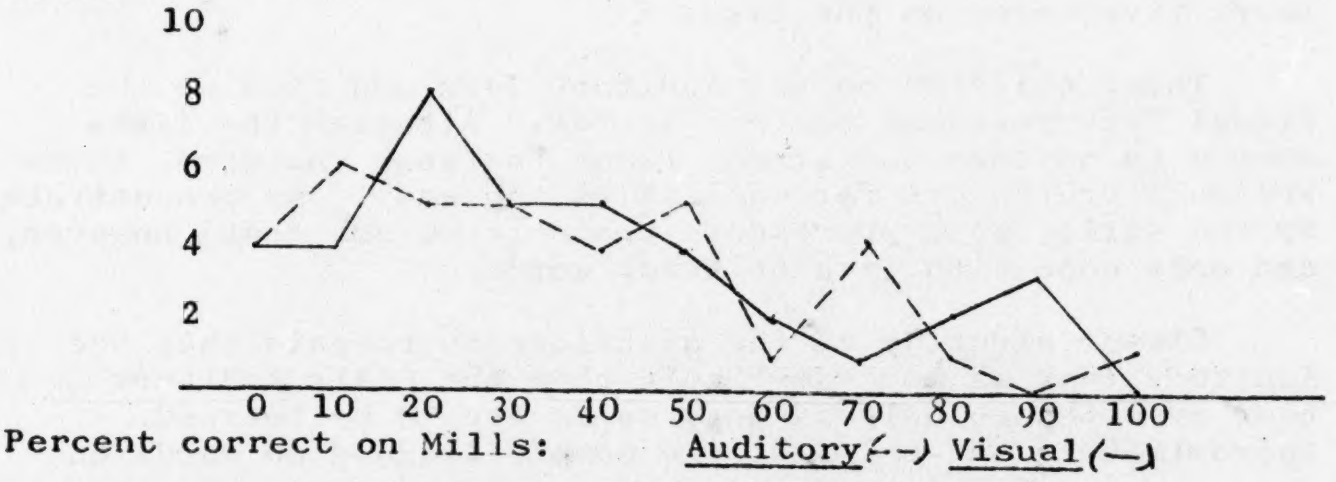
Closer scrutiny of the distribution reveals that the Auditory Task is more difficult than the Mills Auditory test even though half as many words are to be learned. Approximately one-third of the sample learned no words on

FIGURE 4

DISTRIBUTION, MEANS AND STANDARD DEVIATIONS FOR AUDITORY AND VISUAL SCORES ON THE TASKS AND THE MILLS



	Auditory	Visual
Mean:	3.05	4.82
Standard Deviation:	3.09	2.86
Correlation:	.831	



	Auditory	Visual
Mean:	6.24	6.71
Standard Deviation:	4.87	5.26
Correlation:	.586	

the Auditory Task as indicated by the naming score. This was three times the proportion failing to learn on the Mills Auditory test. Whereas the means on the Mills Auditory and Visual tests are similar ($t = 1.41$, $p > .05$), the difference between means on the Auditory and Visual Tasks is statistically significant ($t = 7.58$, $p < .001$).

Sources of difficulty on the Auditory Task could lie in the types of operations expected. Concentration is first on letters in isolation, with auditory discrimination of their sounds, association of sounds with the letters, and finally blending the sounds of letters in sequence. Focus is on the initial letter and word ending rather than on the whole word form. In contrast, words are presented as whole in the Mills Auditory method. Given this presentation the child has greater opportunity to use visual cues as well as letter-sound cues.

Another source of difficulty could stem from the testing conditions. The Tasks were administered to a group of six children whereas the Mills was administered individually. Previous evidence presented when individual and group administrations of the Tasks were discussed, however, shows little effect from group size. Nevertheless, the Mills' directions for administration are considerably less structured than those for the Tasks and enable greater adjustment to the child's learning needs in an individual setting. On the basis of this evidence the Auditory Task seems to be less effective than the Mills. However, such conclusions should remain tentative until evidence concerning predictive validity is presented.

Correlation coefficients among word naming scores on the Mills and Tasks and the total learning-retention scores on the Tasks are shown in Table 11. As can be seen less than 35 per cent of the variation is common to both tests when either the Tasks scores of 10 or 35 are compared with the Mills scores. Although both tests claim to measure word learning, the degree of relationship between them appears to be moderate to low. As previously reported reliability coefficients for the Mills Auditory and Visual Tests for the present sample ($N = 38$) were .850 and .802 respectively. These estimates were lower than reliability coefficients reported by Mills (.969 and .970). Nevertheless, the coefficients were sufficiently high so as to permit a fair test of concurrent validity.

Table 12 contains coefficients showing the relative effectiveness of the Mills and the Tasks in predicting reading. The score used for the Tasks is based on 35 responses. Measures of first grade reading are scores from the Wide Range Achievement Test (WRAT) administered in December and May, and the Vocabulary and Comprehension

TABLE 11

CORRELATION COEFFICIENTS FOR AUDITORY AND VISUAL MILLS SCORES AND TASK SCORES OF 10 AND 35 (N = 38)

	Task Score (=10)		Task Score (=35)	
Mills Score (=20)	Auditory	Visual	Auditory	Visual
Auditory	.607	.493	.594	.541
Visual	.421	.422	.434	.483

$r > .418$; $p < .01$

scores from the Gates-MacGinitie Reading Tests, Primary A, Form 1, administered in May (Gates-Vocabulary and Gates-Comprehension).

Most coefficients shown in Table 12, though statistically significant account for less than 30 per cent of later reading score variation. The Mills and the Tasks are comparable in their prediction of May reading scores. In contrast, the Tasks are more effective than the Mills in predicting December reading.

The validity of the Tasks was examined by comparison of it with the Mills. When word learning on the two tests was compared, an insubstantial portion of the variation was shared by the two tests. If the Mills test is used as a concurrent criterion of validity, then the Tasks are inefficient in measuring this type of word learning. Nevertheless, evidence was presented showing that the Mills was ineffective in predicting later reading. The Auditory Task was similar to the Mills Auditory test in predictive effectiveness. The Visual Task was somewhat more effective than the Mills Visual test, especially in predicting reading in December. Nevertheless, all coefficients were so low as not to justify the Tasks' use for prediction of later reading for this subgroup.

Coefficients of the Tasks with later reading for the Mills - Tasks sample (N = 38), the sample having other readiness measures (N = 108), and the total sample (N = 181) are shown in Table 13. As can be seen, predictive effectiveness improves considerably with sample size. The sample of 38 subjects appears to have been too small to assess the

TABLE 12

CORRELATION COEFFICIENTS FOR AUDITORY AND VISUAL SCORES FOR THE MILLS AND TASKS AND READING SKILLS (N = 38)

	Mills (score=20)		Task (score=35)	
	Auditory	Visual	Auditory	Visual
December Tests:				
Words Learned	.179	.333	.257	.358
WRAT	.386	.287	.434	.546
May Tests:				
Words Learned	.354	.313	.439	.591
WRAT	.460	.408	.460	.493
Phonics	.334	.414	.427	.528
Gates - Vocab.	.519	.389	.455	.473
Gates - Comp.	.462	.458	.462	.554
	$r > .349$; $p < .05$		$r > .445$; $p < .01$	

predictive effectiveness of the Mills and the Tasks fairly. All that can be concluded is that the Tasks are similar or slightly more effective than the Mills in predicting the later reading achievement of pre-readers.

Tasks-Word Learning during first grade. Another estimate of word learning ability was compared with learning measured by the Tasks. Table 14 shows the correlation coefficients between the Tasks, Words Learned-December and Efficiency for the children tested in May and September on the Tasks. Contrary to expectations, correlations are higher for the sample tested in May than for those tested in September. Little more than 40 per cent of the variation is common between the Tasks and Words Learned - December and even less between the Tasks and Efficiency of word learning.

TABLE 13

CORRELATION COEFFICIENTS FOR THE TASK LEARNING-RETENTION SCORES AND READING SCORES FOR THREE SAMPLE SIZES

	Learning-Retention scores on Combined Task		
	N = 38	N = 108	N = 181
December			
Words Learned	.321	.606	.517
Efficiency	.175	.452	.363
WRAT	.513	.613	.602
May			
Words Learned	.538	.597	.623
Efficiency	.360	.459	.462
WRAT	.506	.617	.628
Gates - Voc.	.491	.648	.628
Gates - Comp.	.534	.655	.644
Phonics	.501	.599	.575
	$r > .418$; $p < .01$	$r > .254$; $p < .01$	$r > .254$; $p < .01$
	$r > .325$; $p < .05$	$r > .195$; $p < .05$	$r > .195$; $p < .05$

TABLE 14

CORRELATION COEFFICIENTS FOR TASK AND WORDS LEARNED, DECEMBER, AND EFFICIENCY SCORES FOR MAY AND SEPTEMBER SAMPLES

	Learning-Retention Scores, Combined Task	
Tasks Administered:	May 1969 (N = 71)	September 1969 (N = 110)
Reading Scores December 1969		
Words Learned	.643	.601
Efficiency	.485	.348
	$r > .302$; $p < .01$	$r > .254$; $p < .01$

The samples of learning obtained using the Tasks were found to be highly stable. Nevertheless, the samples were obtained over a short period of time, using special materials not necessarily similar to the child's regular reading materials, by a person other than his regular teacher. The findings show that learning measured under such conditions is not highly similar to the word learning that a child does during the first three months of first grade. The coefficients of correlation, though statistically significant, are not practically significant.

Table 15 shows coefficients between words learned during the first three months of first grade (Words Learned, December), Efficiency of word learning, the Tasks, and May reading scores. Words Learned - December is similar to the Tasks in its usefulness in predicting reading in May as measured by the WRAT and the Gates-MacGinitie tests. The Tasks seems somewhat more effective in predicting phonics than does Words Learned-December. (The Auditory Task is no more effective in predicting later skill in phonics than is the Visual Task.) Whereas the efficiency rating is a good predictor for the May sample, it is quite ineffective predicting reading for the September sample.

The validity of the Tasks as assessed by comparison of it with other word learning measures is of moderate degree. The median validity coefficient between the Mills and the Tasks was .541. Similarly, the degree of relation between the Tasks and word learning during the first three months

TABLE 15

CORRELATION COEFFICIENTS FOR COMBINED TASKS, DECEMBER WORDS LEARNED, EFFICIENCY, AND MAY READING SKILL FOR MAY AND SEPTEMBER SAMPLES

	May Sample (N = 71)			September Sample (N = 110)		
May Reading Scores	Learning-Retention Tasks	Words Learned December	Effic. December	Learning-Retention Tasks	Words Learned December	Effic. December
Phonics	.630	.515	.542	.533	.422	.172
Words Learned May	.534	.605	.529	.680	.738	.419
WRAT	.639	.638	.664	.624	.637	.317
Gates - Voc.	.657	.733	.729	.623	.649	.434
Gates - Comp.	.687	.675	.660	.650	.685	.386
	$r > .302$; $p < .01$.			$r > .254$; $p < .01$ $r > .145$; $p < .05$		

of first grade (Words Learned, December) is moderate: .658 for the May sample and .609 for the September sample.

The Tasks are similar to the other word learning samples (Mills and Words Learned December) for predicting reading achievement in May. The degree of relation between the Tasks and Words Learned- December is similar to that found either between the Tasks and May reading scores or between Words Learned-December and May reading scores.

The finding that words learned during the first three months of first grade (Words Learned, December) showed only a moderate degree of relationship with May reading was unexpected. Correlation coefficients between reading measured at one time and several years later are usually considerably higher, in the range of .67[1] to .82[2]. It may be that standardized measures are better predictors because they contain a wider range of items (untaught as well as taught items) than do samples of words selected from the child's reader. This hypothesis is supported by comparing Words Learned- December and the WRAT (administered in December) in their ability to predict May reading (Table 16).

---

[1]Gertrude H. Hildreth, Results of repeated measurement of pupil achievement, Journal of Educational Psychology, 1930, 21, 286-296.

[2]Arthur E. Traxler, Reading Growth of Secondary School Pupils during a five year period, Educational Records Bulletin, 1950, 54, 96-107.

TABLE 16

CORRELATION COEFFICIENTS FOR DECEMBER WORDS LEARNED, WRAT MAY READING SCORES (N = 131)

May Reading Scores	Words Learned, December	WRAT December
Phonics	.394	.539
Words Learned	.569	.683
WRAT	.567	.768
Gates - Voc.	.645	.712
Gates - Comp.	.649	.714

$r > .254$; $p < .01$

Discussion and Further Observations

It was reported earlier that the Tasks and Words Learned December were similar in predicting May reading. The Efficiency rating was a good predictor for the May sample, but by contrast, quite ineffective for the September sample. The degree of relationship between Words Learned-December and Efficiency is similar for the May and September groups (.647 and .619, respectively). Nevertheless, the mean level of efficiency of the September group was significantly lower than that of the May group in December (62 per cent versus 77 per cent).

The efficiency rating may be a poor predictor to the extent that it reflects the teacher's pacing rather than the child's learning. Further information concerning Tasks and December and May reading scores for subgroups of subjects is shown in Table 17. A comparison of groups for the December Efficiency scores shows that the September sample was generally less efficient in word mastery than was the May sample. In addition, the September sample with the exception of urban boys was introduced to fewer words than was the May sample. In terms of learning on the Tasks, the smaller number of words presented was appropriate for the urban girls but not for the suburban sample.

TABLE 17

MEAN SCORES FOR TASKS AND READING TESTS FOR SUBJECTS GROUPED ACCORDING TO SEX, AREA, AND TASK TESTING TIME

	Task admin. in May				Task admin. in September			
	Urban		Suburban		Urban		Suburban	
	M	F	M	F	M	F	M	F
Number of subjects:	18	15	20	18	21	22	34	33
Tasks:								
Auditory	12	20	15	20	19	12	16	22
Visual	15	20	16	22	21	18	20	24
December reading:								
Words Presented	39	65	51	71	48	39	34	41
Words Learned	23	53	37	67	32	23	21	26
Efficiency	59%	82%	73%	94%	66%	59%	61%	63%
WRAT	26	31	28	32	28	27	27	29
May Reading:								
Words Presented	197	334	193	202	250	199	225	252
Words Learned	122	246	158	178	195	147	171	207
Efficiency	62%	79%	80%	88%	78%	74%	76%	82%
WRAT	34	42	39	42	38	36	39	43
Phonics	14	21	22	24	19	12	26	30
Gates - Vocab.	25	36	34	40	33	27	31	37
Gates - Comp.	14	22	21	24	18	15	16	21

The May and September suburban groups were similar to each other on the Tasks if allowance is made for increase during the summer on the Visual Task. They differed considerably, however, on the number of words presented by December and by May of first grade and in the number of words learned. The September suburban group exceeded the May group in reading nonsense words (phonics); they were similar in pronouncing real words in isolation (WRAT); and the May group exceeded the September group reading words with picture and sentence context (Gates-MacGinitie).

Observation and discussion with the teachers of the suburban children tested in September revealed that one-fourth of the children were being taught by a highly phonetic method and the other children in the sample were exposed to consonant sounds during the first months of instruction. The emphasis on letter-sound associations was less or non-existant in the other suburban area and the urban areas during the first three months of first grade.

Given the variation in reading scores of the September sample, it is easier to understand why prediction using the learning measures was more successful for children in the May sample. The pattern of reading acquisition reflected the type of instructional program to which children were exposed as well as the length and amount of instruction.

Two groups of urban children (girls tested in September and boys tested in May) were least successful in initial word learning. The number of words presented was restricted but the learning efficiency nevertheless was low. The two groups continued to perform similarly on later reading measures.

Children were questioned informally as retention was checked on the Tasks and during the December and May testing. They were asked such questions as "How did you know that word?; Can you think of any word that looks like this word? Can you figure out this word?"

On the word naming section of the retention check of the Tasks, if a child was unsure of a word he usually guessed one of the five words on the Task or less frequently, a word taught on the prior Task. The guess was made with little regard for word form or letter similarity.

Children during the December testing tended not to guess. However, when they did guess, it was usually a word to which they had previously been exposed. Miscues at this stage tended to have some visual similarity to the printed

word -- such as look for book, little for letter, and dog for big. For those children who had been exposed to letter-sound associations, the beginning sound or several unblended sounds in sequence was often given.

Children who made below average progress in May responded similarly to some of the better readers tested in December. Above-average readers in May, guessed at unfamiliar words with a high degree of accuracy. Initial consonants were rarely wrong. Some children could explain what they were doing. "Well, I just sounded it out." "Well, it ends like tip so it must be lip." Most children giving explanations could also pronounce most of the nonsense syllables. Some children seemed to be aware for the first time that by sound-blending or by substituting an initial consonant to a familiar word ending, they could say words. Their performance was usually poor on the first few nonsense items, followed by apparent enjoyment and adequate performance. The better readers could cope most successfully with three-letter sequences. Even though they could read real words of two or fewer syllables, most tended to have difficulty with four-letter nonsense sequences ending with e. The phonics test, an artificial reading situation since word meaning and sentence context cues were removed, was the most frustrating task for most children.

Both the changes in the types of errors made by children in word recognition at times during the first year of reading and the moderate relationships between word learning measured prior to reading instruction, after three months in first grade, and at the end of first grade suggest that the nature of reading changes during the acquisition phase.

More systematic examination of the processes used by children as they acquire reading skill is needed. Examination should provide insight into the reasons why it is difficult to predict later reading achievement from earlier learning or readiness performance.

Observations Concerning the Tasks

In addition to yielding learning and retention scores, the Tasks provided an opportunity for the teacher-researcher to observe the child while learning. Accurate observations were more easily made, of course, when the child was tested individually rather than in a group with five other children.

On the Auditory Task it was possible to observe whether the child could hear differences among initial word sounds (auditory discrimination), if he could successfully associate a sound with a letter (letter-sound association), and using

an initial sound substitution approach, if he could blend an initial sound to a word ending (sound blending).

Similarly on the Visual Task, the teacher-researcher could observe if the child was able to associate a printed form with a pictured meaning, if he could visually select which of the five word forms was the same as a key word held next to the five words (visual discrimination and matching) and if he could identify similarities and differences among the forms of the five words (identification of word form features).

It was possible to see how soon a child learned a word, for example, on the first trial or after several review trials. From the errors that the child made, it was possible to determine how well he understood the Task demands. On the Visual Task, most children who made errors on the naming test selected a word from among the five words taught. There were a few children, however, who selected words from a prior Task or used words not included on the Tasks. Though all of these children would receive no credit for this response, the nature of the response was useful to the teacher-researcher diagnostically to judge the degree of readiness for coping with words.

Several teacher-researchers observed that the poorer learners had more difficulty on the Auditory than on the Visual Task and more difficulty in an individual than a group setting where they could observe the responses of other children. The better learners appeared to do equally well on the two tasks, but preferred the Auditory task when they were successful with it.

The correct-incorrect scoring procedure seemed to mask much additional observational information that was obtained by the teacher-researcher. Further research is needed to determine ways in which these observations can be recorded systematically during the administration of the Task.

Summary

The Word Learning Tasks were found to measure a stable characteristic, and on that basis were judged to be reliable. The learning-retention score shows a high degree of reliability for the Combined Tasks (.929), and an acceptable degree for the Auditory Tasks (.893) and Visual Tasks (.847) separately.

The tasks show a moderate degree of relationship with other measures of word learning (Mills and December- Words Learned) and with May reading. The Tasks were similar to

the other two word learning measures in ability to predict May reading accurately.

On the basis of these comparisons, the Tasks are judged to be reliable and at least as valid as other word learning samples. Given these standardization results, the Tasks are judged to provide a representative sample of word learning to be compared with readiness measures in the next section of the report.

III.

COMPARISON OF THE PREDICTIVE EFFECTIVENESS OF LEARNING AND READINESS MEASURES

Purpose

In the first section of the report, the argument was presented that indicators of readiness measure reading achievement poorly because they assess separate component skills rather than the integrated processes involved in word recognition. It is expected, accordingly, that word learning measures will be significantly more effective in predicting first grade reading than are readiness instruments that measure component skills separately.

To test this expectation, Word Learning Tasks were developed and shown to measure word learning as reliably and validly as other measures of word learning. Readiness measures were selected and compared with the learning measures in terms of their effectiveness in predicting first grade reading achievement.

Methodological Issues

Several methodological issues arose in making this comparison. One concerned the nature of the readiness tests selected, a second, the meaning of prediction, and a third the length of the intervening time span between prediction and achievement measurements.

Selection of Readiness Tests

Readiness measures were selected to represent the best predictors available. One was selected that showed predictive validity coefficients similar to other well-designed readiness tests and measured not only auditory and visual perceptual skills but letter naming (which has been found to be the single best predictor), and knowledge of word meanings. This was the Metropolitan Readiness Test.[1]

---

[1] G. H. Hildreth et.al., Metropolitan Readiness Tests, Manual of Directions. New York: Harcourt Brace & World Inc., 1965.

A second estimate of readiness has been shown in some studies to show a high degree of relationship with reading achievement, namely, kindergarten teacher ratings.[1] These ratings appear to be a composite of a child's performance on informal reading related tasks (e.g., writing letters; knowing the alphabet) and observation of learning and response to new situations in kindergarten. In order to determine if kindergarten teachers' ratings improve in predictive accuracy when specific characteristics upon which to rate the child are included, both a structured and an unstructured rating were obtained for some children in the study. The best rating was then selected empirically.

Prediction of Reading

In many studies, the operational definition of prediction is specified by the statistical procedures used to organize the data. Through the use of correlational techniques, to predict reading operationally means to identify the rank of children in later reading achievement. This definition of predictive effectiveness is frequently used in judging the usefulness of readiness instruments. One original purpose of such tests, however, was to predict which children would fail to make adequate progress in reading. Failure to make adequate progress is usually judged by the failure to achieve a minimal level of skill after a specified interval of instruction. Assessment of readiness using this definition would entail the use of a criterion score of the lowest reading level acceptable. In the evaluation of the predictive effectiveness of learning and readiness tests, I will use both procedures: a correlational analysis to indicate accuracy in ranking and criterion scores to indicate adequate progress.

Time Interval between Prediction and Achievement Measures

Readiness and learning measures were used to predict reading at the end of the first grade. Selection of this time to test reading achievement was arbitrary. For most subjects, this measurement represents eight months of

---

[1] J. A. Ebbesen, Kindergarten teacher ratings as predictors of academic achievement in the primary grades. Journal of Educational Measurement, 1968, 5, 259-64.

reading instruction. Nevertheless, reading instruction was delayed slightly for approximately ten per cent of the subjects, and approximately three per cent of the subjects had not begun reading by the time of the December testing. Thus, although reading achievement testing in May represents eight months of reading instruction for most children, some children receive less.

For most children instructional contingencies were similar in time span but not necessarily in quality or type of instruction or in time spent reading per day. Measurement of reading at the end of first grade, commonly used in prediction studies, is an index only of time span of instruction, not of other relevant dimensions.

The organization of findings on the basis of subgroups controls to some extent for variation in type of instruction and amount of time spent reading daily, but not for quality of instruction. Even if these factors were controlled, it would be impossible to control for home-environmental conditions. Reading achievement, as used in this study, is recognized to reflect other influences as well as instructional effects.

The expectation that readiness and learning tests should predict achievement after one year of instruction may be inappropriate, an unfair test of validity. If designed to measure readiness, how can the tests be expected to take account of all instructional contingencies exerting influence throughout the school year? A fairer test of the usefulness of readiness tests would be their efficacy in predicting initial reading.

Methods

In order to compare the predictive effectiveness of readiness and learning measures over an eight month interval, readiness measures were collected for subjects possessing Task data and follow-up reading data. The following comparisons were made:

1) Word Learning Task scores, words learned after three months of instruction (Words Learned, December). Metropolitan Readiness scores, and Teacher Ratings were compared on their effectiveness to predict reading in May of the first grade.

2) Word Learning Task scores and Metropolitan readiness scores were compared on their effectiveness to predict reading at various times in first grade.

3) Task scores, Words Learned-December, Metropolitan Readiness scores, and Teacher Ratings were compared on their effectiveness to identify children achieving below a selected criterion score.

Instruments

Word Learning Measures

All subjects received one group administration of the Auditory and Visual Tasks. A sample of words learned during the first three months of first grade (Words Learned, December) was also obtained for each subject. (For further description of the Tasks and Words Learned-December see pages 5 to 7 and Appendix A.)

Readiness Measures

The Metropolitan Readiness Tests[1] was selected as an adequately constructed and standardized instrument, similar to other readiness tests in validity and reliability. It shows a strong relationship with other frequently used readiness tests: Pintner-Cunningham, Primary, .76, and Murphy-Durrell Analysis, .80. Reliability coefficients for the total test range from .91 to .94. Coefficients showing the degree of relationship between the Metropolitan readiness test administered in October of Grade 1 with the Reading section of the Metropolitan Achievement Test administered in May of Grade 1 range from .54 to .73 with a median of .65.

Kindergarten teacher ratings of reading readiness are frequently among the most successful predictors of later learning.[2] These ratings may be improved when the teacher is given some guidelines concerning the characteristics to be rated. In this study, kindergarten teachers whose 70 students were tested in May, rated their students using two procedures:

---

[1]Hildreth, loc.cit.

[2]Ebbesen, loc.cit.

1) rating the children on a scale provided by the investigator. Children were rated on seven characteristics frequently reported as related to reading success. (A copy of the rating scale is contained in Appendix C.)

2) classifying the children in one of the four following categories: (1) definitely not ready for reading instruction; (2) possibly ready but will require special help from the teacher; (3) probably ready; (4) ready for reading, will show excellent progress.

The two teacher ratings were highly correlated (.735). Coefficients of correlation between these ratings and concurrent measures of readiness and later reading scores are shown in Table 18. Though similar in predictive effectiveness, the unstructured teacher rating consistently shows higher correlation with later reading than does the structured rating. Because the unstructured teacher rating seemed more useful (contrary to experimenter expectations), only this rating was used in further comparisons. Since the unstructured ratings were collected at the same time as the structered, it is possible that one rating influenced the other.

TABLE 18

CORRELATION COEFFICIENTS OF UNSTRUCTURED TEACHER RATINGS, STRUCTURED TEACHER RATINGS AND METROPOLITAN, TASKS, DECEMBER/WRAT, MAY/WRAT, GATES-MACGINITIE VOCABULARY, AND GATES-MACGINITIE COMPREHENSION. (N = 79)

	Structured Teacher Ratings	Unstructured Teacher Ratings
Structured Teacher Ratings		.735
Metropolitan	.617	.634
Task	.560	.579
WRAT - December	.600	.630
WRAT - May	.541	.584
Gates-MacGinitie - Vocab.	.620	.659
Gates-MacGinitie - Comp.	.641	.653

Unstructured teacher ratings were collected directly from the kindergarten teacher for the May sample. For children tested in September of the first grade, kindergarten teacher ratings were obtained from records available from either the first grade teacher or the school office. The readiness ratings for the September sample were made on either a three- or four-point scale ranging from unready to very ready for reading. Since teachers using the three-point scale marked children approximately one-fourth of the time between the middle and highest rating, this position was regarded as an additional point on the three-point scale, making the rating more comparable to those teachers who had marked on a four-point scale.

Reading Tests

Reading was measured in December after three months in first grade and in May of first grade. In December, in addition to the test of words learned, Words Learned-December, (discussed under Word Learning Measures), the Wide Range Achievement Tests, Reading Section (WRAT) was administered. In May, a sample of words taught was tested (Words Learned, May), the WRAT was readministered, and the Gates-MacGinitie Reading Tests, Primary A, Form 1, and an informal phonics test were given. (For more detailed description of these tests, see pages 7 to 8 and Appendix C.)

Subjects

Teacher Ratings and Metropolitan Reading Readiness scores were obtained for 70 children tested on the Tasks in May and 38 tested in September. Whereas the May sample included 32 urban and 38 suburban subjects, the September sample included only urban subjects. Teacher Ratings (but no Metropolitan scores) were available for 63 suburban subjects tested in September. Table 19 shows the number of children in each group (area and time of testing) for whom complete Task and reading scores were available, and the date of testing on the Metropolitan, and Teacher Ratings. For a more complete description of subject characteristics, see pages 8 to 10.

Findings

Prediction of May Reading: Correlational Analysis

Intercorrelations between the learning and readiness measures, Tasks, Words Learned--December, Metropolitan, and

TABLE 19

NUMBER OF SUBJECTS AND DATE OF ADMINISTRATION OF TASKS, METROPOLITAN, AND TEACHER RATINGS

Tasks Administered:	May				September			
	Urban		Suburban		Urban		Suburban	
	Date	N	Date	N	Date	N	Date	N
Number of subjects for whom Task and Reading Scores were obtained	5/69	33	5/69	38	9/69	43	9/69	67
Date and number of subjects for whom Metropolitan Scores were obtained	9/69	32	5/69	38	9/69	43		0
Number of Subjects with Kindergarten Teacher Ratings	5/69	32	5/69	38	6/69	38*	6/69	63*
Total Number Involved in Prediction Study		32		38		38		63

*Teacher ratings were unavailable for children who had not attended Kindergarten.

Teacher Ratings, are shown in Table 20. The Tasks, Words Learned--December, Metropolitan show fairly high degrees of correlation for all groups tested. The relationship between the Teacher Ratings and the other measures appears to be more complex: for the Kindergarten sample tested on the Tasks in May, between 25 and 55 per cent of the variation is common between Teacher Ratings and other measures. By contrast for the sample tested on the Tasks in September between 3 and 21 per cent common variation is found.

Urban, Suburban Comparisons. Means and standard deviations are shown for four groups of subjects in Table 21. Urban and suburban subjects tested in May show a consistent pattern on the measures. By contrast the samples tested in September show an erratic pattern of mean scores. The

TABLE 20

CORRELATION COEFFICIENTS BETWEEN TASKS, WORDS LEARNED-DECEMBER, METROPOLITAN, AND TEACHER RATINGS FOR URBAN AND SUBURBAN SAMPLES TESTED ON THE TASKS IN MAY AND SEPTEMBER

	Subjects Tested in May on Tasks		Subjects Tested in September on Tasks	
	Urban	Suburban	Urban	Suburban
N	32	38	38	63
Tasks X Words Learned-December	.725	.575	.672	.635
Tasks X Metropolitan	.689	.720	.601	--
Words Learned-December X Metropolitan	.505	.603	.676	--
Tasks X Teacher Ratings	.611	.549	.217	.338
Words Learned-December X Teacher Ratings	.708	.520	.187	.349
Metropolitan X Teacher Ratings	.564	.739	.458	--

September urban sample should be similar to the May urban sample; instead the Teacher Ratings are higher on the average, and the Metropolitan and December Words Learned are lower. The September suburban group shows the highest score on the Tasks and yet receives the lowest Teacher Ratings.

Higher coefficients were found between the May Teacher Ratings and other measures than between the September Teacher Ratings and other measures. This difference in rating effectiveness shown for the May sample may reflect a more valid and reliable rating procedure used by these kindergarten teachers. Because they used the more structured rating scale at the same time as the unstructured rating and because of contact with the teacher-researcher, they may have become more highly motivated and sensitive raters. By contrast, there was no contact between teacher-researchers and kindergarten teachers of the Fall sample. Ratings on a three- or four-point scale were obtained from the child's record. The data collection procedures may be responsible for the pattern of correlation coefficients and mean scores.

TABLE 21

MEANS AND STANDARD DEVIATIONS ON TASKS, WORDS LEARNED-DECEMBER, METROPOLITAN, AND TEACHER RATINGS FOR URBAN AND SUBURBAN SAMPLES TESTED ON THE TASKS IN MAY AND SEPTEMBER

	Subjects Tested in May on Tasks				Subjects Tested in September on Tasks			
	Urban		Suburban		Urban		Suburban	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Number of subjects	32		38		38		63	
Tasks	32.34	18.30	36.13	19.94	35.32	18.27	40.33	19.16
Words Learned-Dec.	36.09	35.84	51.21	34.57	26.61	20.33	23.43	16.09
Metropolitan	60.84	13.80	65.47	12.55	51.05	14.56	--	--
Teacher Ratings[1]	1.94	1.01	2.11	.94	2.24	.63	1.54	1.10

[1]For Teachers' Ratings, 4 indicates that a child is ready for reading and 1 that he appears unready.

The learning and readiness measures were compared for their effectiveness to predict the May reading scores. Table 22 show estimates of the relationship between the learning and readiness measures and reading for the samples tested on the Tasks in May and in September. The May Words Learned represents the child's reading of words that he had learned during the first grade; the WRAT entails reading both familiar and unfamiliar words; the phonics test taps the child's reading of unfamiliar nonsense syllables. The Gates-MacGinitie tests involve reading with picture context (Vocabulary Section) and picture and sentence context (Comprehension Section).

The Metropolitan is most effective in predicting the reading achievement of three of the four samples; approximately 50 per cent of the variation is accounted for. Words Learned-December is the best predictor for Urban subjects tested on the Tasks in September. Generally, however, Words Learned-December and the Tasks are somewhat less effective than the Metropolitan, accounting for approximately 40 per cent of later reading variation. All measures except Teacher Rating appear to be more effective in the prediction of urban than suburban reading. Teacher rating is least effective in prediction in general and in

TABLE 22

CORRELATION COEFFICIENTS BETWEEN MAY READING SCORES AND TASKS, WORDS LEARNED-DECEMBER, METROPOLITAN, AND TEACHER RATINGS FOR THE MAY AND SEPTEMBER SAMPLES

	May Sample							
	Urban (N = 32)				Suburban (N = 38)			
May Reading	Tasks	Words Learned December	Metro-politan	Teacher Ratings	Tasks	Words Learned December	Metro-politan	Teacher Ratings
Words Learned, May	.724	.826	.682	.682	.393	.386	.445	.449
WRAT	.765	.657	.792	.628	.522	.628	.593	.540
Phonics	.657	.523	.704	.495	.599	.471	.486	.465
Gates - Vocabulary	.747	.808	.754	.661	.596	.632	.665	.693
Gates - Comprehension	.730	.676	.746	.605	.673	.650	.675	.727

	September Sample							
	Urban (N = 38)				Suburban (N = 63)			
May Reading								
Words Learned, May	.720	.912	.750	.310	.694	.625	--	.401
WRAT	.622	.794	.642	.188	.645	.666	--	.444
Phonics	.580	.608	.472	.228	.517	.487	--	.340
Gates - Vocabulary	.677	.829	.690	.292	.604	.602	--	.561
Gates - Comprehension	.644	.829	.743	.231	.670	.634	--	.416

particular for the September sample. The prediction was least effective for the task which least resembles normal reading (the phonics tests), and generally most effective for the two standardized tests of reading, the Gates-MacGinitie tests, and the WRAT.

Sex Differences. Because the predictive effectiveness of the four measures was somewhat different for boys and girls, coefficients between these measures and reading in May are shown in Table 23. The pattern of coefficients for girls shown in Table 23 is similar to that for the May sample. When 32 suburban girls tested in September (for whom there were Teacher Ratings but no Metropolitan Scores) are included in the analysis, the coefficient is increased.

There is a tendency for Tasks and the Metropolitan to be slightly more effective than Words Learned-December in predicting May reading, whereas Words Learned-December is somewhat more effective than the other measures for boys. When 31 suburban boys tested in September (for whom there were Teacher Ratings but no Metropolitan scores) are included in the analysis, the coefficients show a marked decrease. Apparently, teachers are least accurate in rating boys in the September samples.

Measures of word learning either in the classroom (Words Learned, December) or under special conditions (Tasks) are not generally more effective on predicting May reading achievement than is a traditional readiness measure, the Metropolitan. There were, however, two groups of children for whom May reading was best predicted by the word learning during the first three months of school (Words Learned, December): the boys and the urban sample tested in September. Nearly half of the children in the latter group were from Spanish-speaking homes.

Prediction of Reading at Selected Time Intervals

This section contains an analysis of data to determine if the effectiveness of readiness measures to predict word learning is negatively affected by the period intervening between the administration of readiness and reading measures. As previously argued, the readiness measures cannot be expected to account for all instructional and social contingencies occurring throughout the period of one year. But if readiness measures are tapping characteristics required for word learning, we would expect to observe a substantial correlation between the readiness test and word learning measured at approximately the same time with decreasing correlation coefficients as instructional factors intervene and developmental changes occur.

TABLE 23

CORRELATION COEFFICIENTS BETWEEN MAY READING SCORES AND TASKS, WORDS LEARNED-DECEMBER, METROPOLITAN, AND TEACHER RATINGS FOR BOYS AND GIRLS

	Boys (N = 57)					Girls (N = 51)				
May Reading	Tasks	Words Learned December	Metro-politan	Teacher Rating		Tasks	Words Learned December	Metro-politan	Teacher Rating	
N	57	57	57	57	88	51	51	51	51	83
Words Learned, May	.570	.663	.566	.440	.308	.610	.623	.657	.473	.364
WRAT	.611	.779	.699	.444	.176	.616	.619	.717	.475	.350
Phonics	.544	.631	.577	.371	.249	.668	.530	.615	.447	.134
Gates - Vocabulary	.608	.785	.661	.521	.373	.686	.757	.779	.531	.442
Gates - Comprehension	.632	.790	.643	.455	.313	.671	.674	.846	.574	.472

Correlation coefficients between readiness measures and word learning assessed prior to first grade instruction (TASKS), after three months (WRAT), and after eight and a half months in first grade (WRAT) were computed. Figure 5a shows graphically the changes in the magnitude of these coefficients. A substantial relationship seems to exist between the Metropolitan and pre-instructional word learning (Tasks) which tends to increase slightly in December, and to increase in May, to a level higher than that existing before instruction. The findings do not confirm the expectations that a decreased relationship exists between readiness and reading measures as the instructional interval increases.

By contrast when coefficients between the Tasks and reading as measured by the WRAT in December and May are plotted (See Figure 5b) the expected pattern occurs: coefficients decrease slightly with time. It appears that readiness instruments become better predictors during first grade whereas the learning measure becomes less effective. Since the readiness test measures skills acquired in the past, its increasing predictive effectiveness may reflect its sensitivity to numerous environmental contingencies that affected the child in the past and are still exerting influence on him. A learning task, to a lesser extent, reflects past experiences. Rather, it taps present learning skill under one set of conditions.

The time interval studied was short; therefore, the trends shown could represent chance fluctuation. Study of the coefficients over a longer period of time is needed before reliable conclusions can be drawn concerning the relationship of readiness and learning measures to reading. The initial trends during the first grade study suggest, however, that readiness measures will maintain a stable relationship with subsequent learning or achievement and by contrast, learning measure will become less effective in prediction with time as motivational and environmental factors influence the learning process.

Minimum Criterion Score Analysis

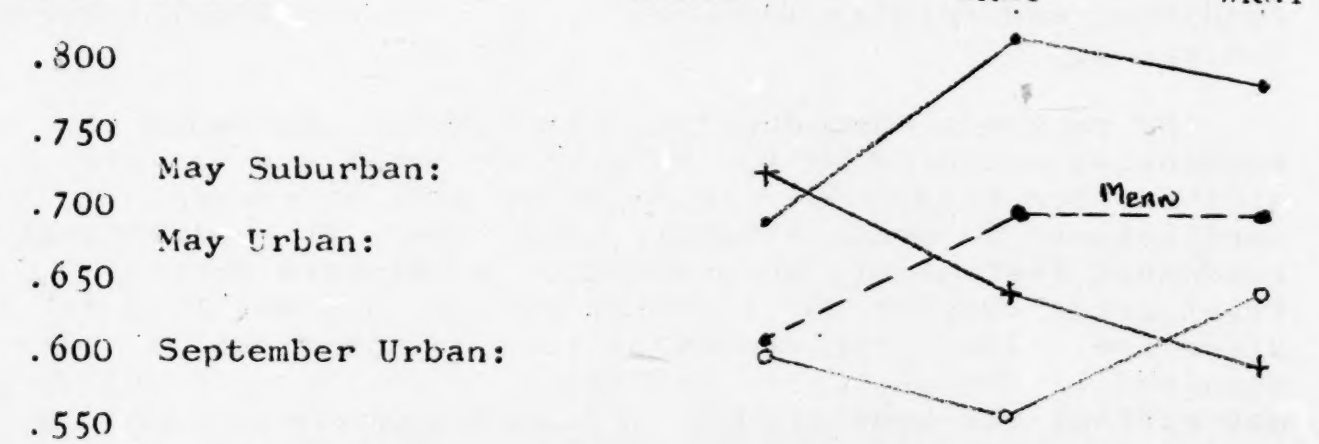
The preceding analyses have shown that the Tasks and Words Learned-December show similar, but slightly less relationship with May reading than does the Metropolitan. When scores were plotted between the Tasks and May reading (WRAT),[1] the spread of the scores made the following

[1]See Appendix E.

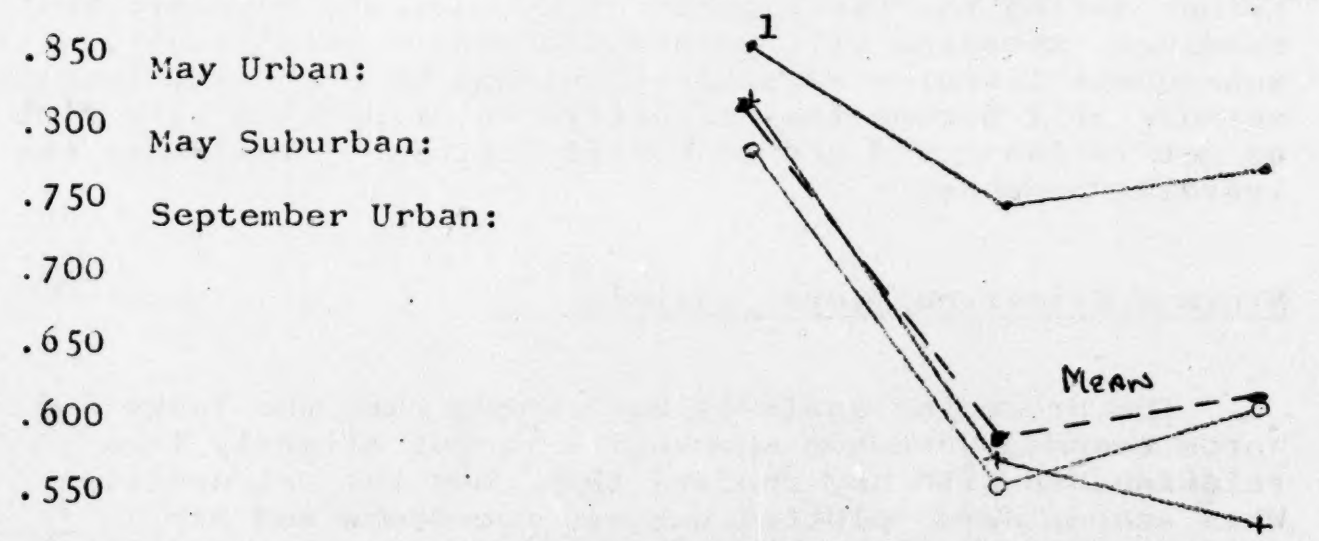
FIGURE 5

TRENDS BASED ON CORRELATION COEFFICIENTS BETWEEN READINESS AND LEARNING MEASURES AND READING ASSESSED DURING FIRST GRADE

a. Readiness vs. Reading (Metropolitan, Tasks, and WRAT)



b. Tasks vs. Reading (Tasks and WRAT)



[1]These points were based on correlations between Auditory and Visual Task scores.

conclusions apparent: if a child did poorly on Tasks, it was difficult to predict how he would achieve in later reading. By contrast if he showed a fifty percent or better performance on the Tasks, chances were very high that he would do well in reading. The same pattern was shown for the Metropolitan: scores in the lower third of the distribution showed a low relationship with reading whereas those in the upper part showed a high relationship. The relationship between Words Learned-December and May reading tended to be consistent throughout the distribution.

The shape of these score distributions is important, especially if the teacher has used these measures to help her identify children who need special help in reading.

This section contains an analysis of data to determine the efficiency of the learning and readiness estimates to identify children who will have extreme difficulty in learning to read.

"Extreme difficulty in learning to read" is defined as the ability to read four or fewer words on the WRAT at the end of first grade. If the child could not read four or more words on the WRAT (e.g., cat, see, red, to, big) by the end of grade one, he was identified as a disabled reader. This cut-off score is equivalent to the average reading performance of children in the fourth month of first grade. Scores for the four learning and readiness measures were available for 108 children. From this sample 16 children (15 per cent) were identified as disabled readers. In Table 24 various cut-off scores are shown for the Task, Words Learned-December, Metropolitan, and Teacher Ratings. Next to the selected scores from the learning and readiness measures are adequate and disabled readers (as identified by the WRAT criterion score). A useful cut-off score on learning and readiness instruments is one that maximizes the number of children correctly identified as not ready for reading (--) or ready for reading (++) and minimizes the number of false negatives: children who were predicted to have difficulty but in fact showed acceptable performance (-+) and the number of false positives: those who were predicted to succeed in reading but were identified as disabled readers at the end of grade 1 (+-).

The information contained in Table 24 can be interpreted as follows: if a teacher used the Tasks to select children needing special instruction, the cut-off score of 14 correct on the Tasks would correctly identify only four children among the 16 needing special instruction and overlook 12. In addition, five children given special instruction would

TABLE 24

SELECTED CUT-OFF SCORES ON THE TASK, METROPOLITAN, TEACHER RATING AND WORDS LEARNED--DECEMBER FOR IDENTIFICATION OF ADEQUATE READERS (30 OR ABOVE ON WRAT--MAY) AND DISABLED READERS (BELOW 30 ON WRAT--MAY) N = 108

Task					Metropolitan					Teacher Rating					Words Learned--December				
Cut-off Score[1]	--	-+	+-	++	Cut-off Score	--	-+	+-	++	Cut-off Score	--	-+	+-	++	Cut-off Score	--	-+	+-	++
14	4	5	12	87	30	6	2	10	90	low	4	1	12	91	0	5	1	11	91
18	10	12	6	80	42	10	8	6	84						3	13	3	3	89
20	13	17	3	75	48	12	17	4	75	med-low	11	14	5	78	6	14	7	2	85
22	15	23	1	69	52	14	25	2	67						8	15	8	1	84
23	16	28	0	64	64	16	51	0	41	med-high	16	51	0	41	26	16	27	0	65

[1]The first of each pair of signs refers to the cut-off score listed. Scores below a particular cut-off score are indicated by -, and scores equal to or above the cut-off score, by +. The second of each pair of signs refers to the criterion score on the WRAT--May; scores below 30 are indicated by - and scores equal to or above 30 by +. For example, in the first row of scores on the Task, four children scored below the cut-off score of 14 and below 30 on the WRAT. Five scored below the cut-off of 14 but equal to or above 30 on the WRAT (False Negatives). Twelve scored equal to or above the cut-off of 14, but below 30 on the WRAT (False Positives). Eighty-seven scored equal to or above the cut-off of 14 and equal to or above 30 on the WRAT.

not have needed it. The most optimal score of 20 on the Tasks would correctly identify 13 children needing special instruction but miss 3. In order to give instruction to this many children needing it, 30 children would be given special instruction; but 17 of these false-negatives would not have needed it. Using the Task score of 20, 88 out of the 108 children (81 per cent) are correctly identified.

A cut-off score of 48 on the Metropolitan yields a selection efficiency similar to a cut-off score of 20 on the Task. A cut-off score of 42 on the Metropolitan correctly identifies 10 of the 16 children needing special instruction. In addition eight children not needing it would receive it.

If a first-grade teacher tried to use the kindergarten Teacher Rating of low or unready for reading as a way of identifying children who needed special instruction, four children would be selected. Most of the children (12 false positives) needing help would be overlooked while a different 12 (false negatives) who did not need special help would receive it.

The measure of learning during the first three months of first grade (Words Learned, December) proves to be the most effective way of selecting children who need special help. Sixteen children in the sample failed to learn three words by December. Thirteen of these were children who at the end of first grade would be identified as disabled readers. Only three would be overlooked, and three not needing instruction would be given it. One hundred and two of the 108 children (94%) would be correctly identified.

If the purpose of readiness measures is to help the first grade teacher identify children who will need special readiness instruction, neither standardized readiness measures, brief samples of word learning nor the ratings based on kindergarten teacher observations are sufficiently accurate. Using the optimal score on each measure, the number of children misidentified exceeds the number of potentially disabled readers correctly identified. Only careful observation and assessment of a child's learning during the first few months of first grade leads to appropriate identification.

Further analysis was undertaken to determine if two or more of the measures obtained before reading instruction in combination would increase their accuracy in selecting disabled readers. Using the criterion scores of 20 on the Tasks, 43 on the Metropolitan, and a rating of 2 (possibly reading for reading), the number of children correctly and incorrectly identified using all possible combinations of tests are shown in Table 25. Children were selected if they performed below the criterion on two tests, on two of three

tests, and on all three tests. As can be seen, the effectiveness of the Tasks to select is increased by using it in combination with either the Metropolitan or Teacher Ratings. The Metropolitan and Teacher Ratings are more effective singly than in combination with each other (using the criteria: number of disabled readers correctly identified in comparison with the number of children incorrectly identified.) When two out of three below criterion scores are used, the pre-instruction measures are more effective than any of the pre-instruction measures singly or in combination, but less effective than Words Learned-December.

TABLE 25

COMBINATIONS OF SCORES ON THE TASK (CUT-OFF SCORE = 20), METROPOLITAN (CUT-OFF SCORE = 48), AND TEACHER RATING (CUT-OFF RATING = MEDIUM LOW) FOR THE IDENTIFICATION OF ADEQUATE READERS (30 OR ABOVE ON WRAT-MAY) AND DISABLED READERS (BELOW 30 ON WRAT--MAY)
N = 108

Combinations of Measures	--[1]	-+	+-	++
Task and Teacher Rating	10	6	6	86
Task and Metropolitan	10	6	6	86
Metropolitan and Teacher Rating	7	6	9	86
Below Cut-off Scores on Two of Three Measures	13	9	3	83
Below Cut-off Scores on All Three Measures	7	3	9	89

[1]These signs can be interpreted similarly to those in Table 24.

A tabular description of urban and suburban children correctly identified, misidentified, and overlooked by each of the learning and readiness measures is shown in Appendix E. The table reveals that 11 of 16 children needing help are urban boys. The group most frequently misidentified by all measures was suburban boys. Those most frequently misidentified by tests, Metropolitan and Tasks, were the urban girls in the September sample. The Teacher Rating for the May

sample was the most effective way to identify urban boys. By contrast, Teacher Rating for the September sample was the least effective identification procedure.

This analysis has been oversimplified by using the same criterion scores for each group of subjects. Nevertheless, the following tentative conclusions can be drawn:

1. Pre-instruction measures singly or in combination are less effective than word learning during the first three months of school (Words Learned, December) in identifying disabled readers.

2. On the basis of findings from the urban sample tested in May versus those from the urban sample tested in September, it is felt that teacher ratings could be made more effective if Kindergarten teachers were given guidelines and incentives for careful ratings.

IV.

SUMMARY

The two main objectives of the study were to develop and evaluate the usefulness of word learning tasks and to compare the effectiveness of the word learning tasks with readiness measures in predicting reading in grade one.

Conclusion: Word Learning Tasks

Development

An auditory and a visual word learning task were developed as miniature versions of the learning that a child might later experience.[1] The directions, procedures, and interpretation of responses were clearly defined so that the Tasks might be used by persons with relatively little testing experience. The Tasks were also designed so that they could be administered to small groups of children as well as to individuals. For clinical purposes, it was hoped that both Tasks could be administered within the period of one day. On the basis of pilot study results, it was judged that the successive administration of the Auditory and Visual Tasks on the same day was overwhelming to most pre-readers. Nevertheless, the success that many children showed with the ten-word sample on the Mills in the main study leaves this question unanswered.

The Tasks were constructed to yield five samples of word learning for each Task. Pilot research findings showed that some children showed an increase in learning with further review, whereas others remained constant or showed a decrease. Similar patterns were found for the subjects in the main study. Increase patterns may represent the ability of the child to profit from review, whereas decrease patterns may reflect the tendency of the child to become confused. Further analysis of the data is needed to determine the nature of the patterns before any conclusions can be drawn concerning their relationship to learning.

Reliability and Validity

The characteristics measured by the Tasks are sufficiently stable over the period of one week to indicate that the Combined

---

[1] See Appendix A.

Task (Auditory and Visual) is reliable (.929). The Auditory and Visual learning-retention scores are somewhat less reliable (.893 and .847, respectively), but well within the range commonly found for learning tasks. Word learning in group and individual settings is similar.

The validity of the Tasks as assessed by their comparison with other word learning measures is of moderate degree. The median validity coefficient between the Mills and the Tasks was .541. Similarly, the degree of relationship between the Tasks and word learning during the first three months of first grade (Words Learned, December) is moderate: .658 for the May sample and .609 for the September sample.

The Tasks are similar to the other word learning samples (Mills and Words Learned, December) for predicting reading achievement in May. The degree of relationship between the Tasks and Words Learned, December is similar to that found between the Tasks and May reading scores and between Words Learned, December and May reading scores.

Auditory and Visual Tasks

Findings from research[1] show that many children can learn to read successfully by an approach which emphasizes either sight words or phonics. However, some children (about 10 per cent) with apparent strength in auditory areas but difficulty in visual ones learn better by a phonics or linguistic method. Likewise, some children with visual strength and auditory difficulty learn to read better through a visual-sight word approach. Other children (approximately 5 per cent) have difficulty learning by any method.

The findings from the Tasks show that 10 per cent of the children showed a 30 per cent or greater score discrepancy in favor of visual word learning, while considerably fewer, 1.7 per cent, showed a similar discrepancy in favor of the auditory method. The disparity among percentages showing a strong visual versus a strong auditory modality may reflect differences between the Tasks or may reveal real differences between the auditory and visual word learning of pre-readers. It was found, however, that the Modality score discrepancies were not stable over the period of one week. Some pre-readers who exhibited low performance on both Tasks may show modality preferences once they begin to acquire reading skills. In other words the tasks may have been so difficult for them that any existing modality preference was not allowed to reveal itself. Although the Tasks may have been insensitive to

---

[1] de Hirsh, loc. cit.

modality preferences, since the discrepancies that did occur during one test period were not replicated one week later, further research is needed to determine the extent to which modality preferences revealed by perceptual tests are stable over time.

The Auditory Task was more difficult to construct because it involved the compression of activities which normally occur over an extended time span into a short one. Although pilot research was undertaken to determine the most appropriate procedures for the Auditory Task,[1] findings in the main study suggest that its format should be restudied. The Auditory Task of five words was found to be more difficult than the Mills Auditory test of ten words. Nevertheless, as previously discussed, the nature of the two auditory tests may be different. Patterns of learning on the four recognition tests following the teaching phase of the Task tended to show a decrease trend in the pilot research. Further analysis of main study data will be made to see if this pattern is replicated. The Auditory Task was constructed to be as similar as possible to later phonics training. By designing the Task in this way, I suspect that children who had already acquired some facility with initial letter sounds could cope with the Task, whereas other children were unable to learn as intended. It is puzzling, however, that the Auditory Task was highly reliable and showed a high degree of correlation with the Visual Task (the two Tasks seem to be measuring essentially the same characteristic). The Auditory Task was similar to, but slightly less effective than, the Visual Task in predicting reading in December and May of first grade.

On the basis of the findings, it is concluded that the Tasks are not useful in identifying modalities through which disabled learners can acquire reading skills. For most children the Tasks appear to measure similar abilities. Further research is needed to specify the relationship between reading acquisition and auditory skill (phoneme discrimination, sementing, blending) before consideration should be given to the redevelopment of an auditory-word learning task.

Observations

The teacher-researchers reported that diagnostically useful information, such as skills employed while the child attempted to learn words and ability to attend to the task, was revealed during the teaching sessions though not reflected in the learning score. Although the teacher-researchers informally recorded the behavior of some children, no formal evaluations were kept that would allow the usefulness of the observational data to be examined. In future research, systematic ratings should be made by the Tasks administrator to supplement and to be compared.

---

[1] See Appendix A.

with the learning score. The Visual Task, and possibly the Auditory Task, appears to provide a useful format for trial teaching during the first months of first grade to see if children have difficulty discriminating, learning, and remembering words. In addition, the Visual Task should be useful in kindergarten to identify children who are able to learn words.

Conclusions: Predictive Effectiveness of the Tasks and Other Measures

One criteria for evaluating the usefulness of the Tasks was that they "should be as effective or more effective than currently available test results in predicting later learning."[1] A comparison of four measures for the prediction of reading in May of first grade revealed that the Metropolitan Readiness Tests was generally the best predictor, although words learned during the first three months of first grade (Words Learned, December) was a better predictor for urban subjects tested in the Fall. The Tasks was consistently (but not significantly) less effective in prediction than the Metropolitan and Words Learned-December; Teacher Rating was generally the least effective for prediction.

It was suggested that the moderate correlation coefficients typically found between readiness and reading might be a function of the time intervals between measures; that is, the greater the time span (and other intervening activities), the lower the correlation. Coefficients between the Metropolitan and reading (WRAT) showed a tendency to increase from December to May in first grade, whereas coefficients between the Tasks and reading (WRAT) displayed a tendency to decrease during the same interval.

The difference in trends may reflect compositional differences between the two tests. The Metropolitan measured prior learning and thus may be sensitive to motivational and socio-familial factors which affected and continue to influence the child. The Tasks, by contrast, measures acquistion of new learning and appears to depend less on past learning.

Further study will be undertaken to see if the trends displayed after a short time interval become more pronounced after two or three years. If they do, in the

---

[1] See proposal for present study, p. 2.

direction predicted, it will be concluded that word learning at one time is useful in predicting word learning after a relatively short interval, but that the predictive effectiveness decreases as the time interval increases and other factors influence learning.

Words Learned-December and the Metropolitan showed similar degrees of correlation with reading in May of first grade. Nevertheless, to identify children with low reading skill at the end of first grade, learning during the first three months of first grade (Words Learned, December) was the most useful measure. No measures obtained prior to reading instruction (Tasks, Metropolitan, Teacher Rating) were as efficient as Words Learned--December in identifying children who would have most difficulty reading. These findings suggest that the most useful information for identifying children in need of special help can be obtained by the first grade teacher during the first few months of first grade. Children who have learned only a few words (3 in this study) tend to be the children who show least progress in reading and who need special help.

Implications for Practice

The Word Learning Tasks are reliable and valid for the purpose of identifying children who will show adequate progress in reading, but are practically ineffective identifying children who will have difficulty reading. Thus, if the goal of the teacher is to select children (for example, Kindergarteners) who will cope successfully with word learning, the Tasks would be an appropriate instrument.

The Visual Task, as constructed, seems more appropriate for pre-readers than does the Auditory Task. Both, however, provide the opportunity to observe a child as he learns words: on the Visual Task, the teacher can observe whether or not the child can match word forms and note differences between them. On the Auditory Task, the teacher can observe whether the child is able to discriminate sounds, to associate a sound with a letter, and to blend sounds. It was noted that on the Auditory Task, performance tended to be bimodal: either children were able to cope (probably because of prior learning in auditory areas) or were unable to cope using auditorily-based skills (these children tended to guess or use visual form cues). It is judged that observation of these skills in the context of word learning is more useful information for the teacher than the testing of component skills with separate tests.

If the teacher's purpose is to select children who will have difficulty learning to read, pre-reading instruction measures are ineffective. Only about half the children identified as potentially disabled will in fact experience difficulty; the others will make adequate progress. The most effective way to identify children in need of special help is to monitor their reading during the first few months of school. In the present study, children were tested in December on a sample of words which had been presented during the first three months of first grade. Those children who learned 3 or less words were the ones who could only read a few words by the end of first grade.

It was concluded that the Tasks are ineffective for identifying disabled readers. Nevertheless, if teachers use the Tasks during the first few months of first grade, relevant diagnostic information might be revealed to aid in the design of appropriate instruction.

If a teacher or researcher wishes to predict reading performance at the end of first grade, a readiness test shows higher correspondence to later achievement than does a learning test. Whereas the correlation coefficients between readiness tests and reading seemed to remain the same or increase at selected time intervals during first grade, coefficients between word learning and reading measures tend to show a slight decline.

Future Research

Both the changes in the types of errors made by children in word recognition at times during the first year of reading and the moderate relationships between word learning measured prior to reading instruction, after three months in first grade, and at the end of first grade suggest that the nature of reading changes during the acquisition phase. More systematic examination of the processes used by children as they acquire reading skill is needed.

Before further revision of the Auditory Task is undertaken, further study of the relationship between perceptual and reading skills is needed. Abstraction of phonemes from words may be a process that can functionally occur only after considerable experience with words. Awareness of letters may provide visual control to help the child to abstract the phoneme from its embedded position in the syllable. Teaching isolated sounds may change the nature of reading acquisition from one of noting similarities and differences and forming generalizations to one of

memorizing associations. The slow rate at which some children responded to a strict phonics approach suggests that they need some experience with word forms before they can abstract and synthesize sounds.

The teacher-researchers reported that diagnostically useful information was revealed during the teaching sessions though not reflected in the learning score. Further research is needed to determine ways in which these observations can be recorded systematically during the administration of the Task.

Further investigation of the measurement of learning or component skill characteristics needs to be conducted within the child's instructional context. The present study showed pre-instructional measures to be relatively ineffective in identifying children who need special help or in specifying the type of instruction needed. Further work is needed on the development of procedures and tasks which can be used as part of the ongoing group instruction. This implies both that teachers need to acquire informal diagnostic skill, and that instructional materials need to include differential tasks.

APPENDIX A

DEVELOPMENT OF THE WORD LEARNING TASKS

AND

TASK MATERIALS AND DIRECTIONS

Development of the Word Learning Tasks

Visual Method.

Development of the Word Learning Tasks entailed the definition of methods, selection of word sample, construction of materials, and standardization of test procedures. Teachers' manuals of five reading series which use a sight-word approach at the primer level were consulted. From this survey, three types of teaching activities were identified: association of the word with the child's prior experience using pictures and discussion; comparison of word forms to note similarities and differences; and observation of such characteristics of the word form as length, initial or ending letters, and configuration.

Word samples for the Visual Task were obtained from five commonly used series: Allyn and Bacon, 1963; Ginn and Co., 100 Edition, 1966; Houghton Mifflin, 1966; Macmillan, 1965; and Scott, Foresman, 1965. The criteria for selection of words to be used included: (a) the words must be used in at least two of the five series; (b) at least two series must agree on the grade-level placement of the word; (c) the word must be easily illustrated; (d) the word must also be on the Dolch list of 95 common nouns. The above criteria resulted in a list of 20 words designated as primer level.

Pretests containing the 20 words were given to approximately 50 children. Five children in this sample knew some words. Words known by several children (e.g., dog) were eliminated from the pool of visual words.

Pilot research was undertaken to determine the optimal number of words to teach to pre-readers. Trial lists of 3, 5, 7 and 9 words were selected and two or three of the list-lengths were taught to nine children. It was found that three words were too few for five of the children. These children performed better on lists of either five or seven than nine words, indicating that the latter number had inhibited learning. The choice between five and seven words was based on observational rather than performance data. The task of five words seemed to provide sufficient ceiling for most pre-readers, but did not overwhelm the less adequate learners. On this basis, it was decided that five words would be taught using the visual procedure and five using the auditory procedure.

Words were assigned to two forms of the visual task, equating the form lists on length and configuration where possible. The final two sets of words selected for the pilot research comparison were:

children, house, tree, girl, bus
rabbit, wagon, farm, duck, boy

Auditory Method

Similarly several phonic or "linguistic" reading series were consulted. The activities identified (auditory discrimination, rhyming, forming of letter-sound concepts, blending and consonant substitution), frequently occurred over an extended time span. Various ways of combining these activities within one learning session were considered. Pilot research involving two possible instructional sequences was undertaken to determine if children could cope with the task sequence and to see which of the two procedures resulted in better learning and/or appropriate behavior.

Seven 4-year-old nursery school children were given the task using procedure A (for each sound, auditory discrimination was checked, the letter's sound was introduced, and application was made to the word through blending). The same children were also given the task using procedure B: (teaching auditory discrimination of all sounds, then the related letter-sound association, and finally application of this learning in words). Four children given procedure A first, obtained a mean score of 9.3 on A and 5.0 on B. Three children given procedure B first, obtained a mean score of 16.5 on A and 12.0 on B. The findings show somewhat better learning under condition A.

The two researchers administering the task reported that with the repetition of the auditory discrimination task under condition B, children seemed to learn the auditory skills. Under procedure A, children were less successful with auditory discrimination and responding correctly to letter-sound associations.

Because procedure B seemed to be more effective in developing auditory skills, it was tentatively accepted for further study even though it appeared to be the more difficult of the two procedures. Further pilot research was undertaken to determine how kindergarten children performed on the task and how the auditory task compared with a visual task using a similar time period.

Selection of words for the auditory approach proved to be somewhat arbitrary. As with the selection of sight words, several common reading series, phonics or linguistic in approach, were consulted. Because of the wide variety of words introduced among the series and the inconsistency regarding sequence of presentation, words had to be selected in a different manner from the sight words. The

researchers noted, however, that almost all the series introduced C-V-C words first, and that the short vowels a and i were introduced first, respectively. Beyond these two similarities, the series appeared to vary considerably.

Using the above information as a guideline, the selection of words was limited further by the following criteria:

a) One V-C word in each set would be the core element (word ending) of the other four words.

b) Words would be selected so that letters similar in sound and configuration would not be included in the same set.

The two sets of words selected to be used with the auditory method were:

at, hat, sat, cat, bat
an, pan, fan, ran, man

Pilot Studies

A series of pilot studies were then undertaken in order to further observe children on the auditory tasks, to compare the auditory and the visual task, to determine if both tasks could be administered on the same day, and to identify the optimal time after learning to check retention. In addition, the pilot research was used to further standardize directions and to explore the possibility of using the task with a group of children.

The third pilot study was undertaken to see if both tasks could be given in one session. Three above-average children and five average learners were given the two tasks consecutively. The first group (above average) received a mean score of 15 on the first administration and 16 on the second. Although they were tired by the end of the second task, they had been able to attend for a period of approximately 50 minutes. The second group (average) received a mean score of 12.8 on the first administration and 8.4 on the second. They were obviously fatigued and frustrated by the task. By contrast a third group, five average learners, were given the tasks on each of two days. Their attention and interest remained high on both days. Their mean score on the first administration was 11.8 and on the second 16.0. The decision was made to administer the tasks on separate days. A comparison of mean scores on the first and second administration suggests that there may be a large practice effect. Perhaps children "learn to learn." Five subjects is, however, too small to yield reliable conclusions.

In order to determine the most useful interval between the learning and the follow-up retentive check, several children were tested 2 hours and either 24 or 48 hours following the learning. Retention measured two hours later appeared to be similar to that in the final learning check. Table 26 shows, as would be expected, that the 24-hour check showed considerably less forgetting than did the 48-hour check.

TABLE 26

PERCENT OF LEARNING AND RETENTION FOR 24 and 48 HOUR INTERVALS.

	24 Hours (N=5)		48 Hours (N=3)	
	Learning	Retention	Learning	Retention
Auditory	65.6	53.0	48.0	27.0
Visual	78.4	68.0	61.2	34.0

Twenty-four hours was selected as a useful interval since it was a sufficient time period for some forgetting to occur and yet easier for data collection. It can be noted in Table 26 that prereaders perform better on the visual than the auditory test. Though the auditory method appears to force the children to depend on auditory skills, it also enables less learning to occur within a given time span.

A somewhat larger group of children was used in the final pilot study to determine if procedures were adequate for group administration and to make further observations concerning the auditory test. Four groups of kindergarten prereaders were selected. Two were given the auditory task first and two the visual. Table 27 shows the mean scores that the four groups received on the auditory and visual tasks.

The procedures seemed to work nearly as well for group as for individual administration. Children responded well both to the auditory and to the visual tasks, even though the learning on the auditory task tended to be approximately two points (of 25) lower than on the visual task. There appeared to be an increase in learning on the second administration but this tendency was not consistent when the visual task was administered first, and the auditory second.

TABLE 27

MEAN LEARNING SCORES (TOTAL POSSIBLE = 25) FOR FOUR GROUPS OF SUBJECTS ON THE AUDITORY AND VISUAL TASKS.

		Auditory	Visual		Visual	Auditory
	N	1 Admin.	2 Admin.	N	1 Admin.	2 Admin.
Group 1	4	11.50	13.25	4	18.00	14.00
Group 2	4	14.75	16.50	5	19.40	20.40

Directions given by the two teacher-researchers were taped and comparison was made to identify inconsistencies with the printed directions and with each other. The directions were revised on the basis of this examination so as to be less ambiguous and more natural. Format of the answer booklets was inspected. Answers for each test were put in long rectangular boxes to help the child keep his place. Each retest sheet was made a different color so that the researcher could more clearly describe the page and check to see that children had followed directions.

The scores from each multiple-choice test for the same four groups of subjects are in Table 28. As can be seen, learning on the Auditory multiple-device tests seems to decrease slightly by the fourth test. By contrast, learning on the Visual multiple-device tests tends to increase with each retest. Half the Auditory and half the Visual tests were administered first. When the pattern of scores on the first administration is compared with that on the second, there is a tendency for children to do better on the first tests of the first administration and then to show a decline and for the opposite to occur on the second administration.

TABLE 23

MEAN SCORES FOR EACH MULTIPLE-CHOICE CHECK (POSSIBLE SCORE =5) FOLLOWING LEARNING ON THE AUDITORY AND VISUAL TASKS AND FOR THE FIRST AND SECOND ADMINISTRATION.

	1	2	3	4	Total Average
Auditory	3.25	2.94	3.13	2.94	2.56
Visual	3.13	3.31	3.50	3.44	3.50
First Admin.	3.38	3.69	3.13	3.00	3.30
Second Admin.	3.00	3.50	3.75	3.50	3.44

Although the evaluation of learning trends was not an objective of the present study, four multiple-choice checks were retained so that the patterns could be studied at some later time. More important, a sufficient sampling of responses were needed to increase chances of total scores being reliable.

In the following section of Appendix A, the task materials and directions as developed through the pilot studies are shown.

Task Materials and Directions

On the following page are samples of the materials and the directions that were developed through pilot research for the Visual Task and Auditory task.

PRE-TEST

Materials:

Pre-test at back of answer booklet.

PRE - TEST Name_____

v1	rabbit	wagon	duck	farm	boy
a1	at	bat	hat	cat	sat
v2	children	house	girl	tree	bus
a2	an	man	fan	pan	ran

Directions:

(Administered to each child individually. Ex. records answers.)

I am trying to find some words that you don't know so that I can teach them to you. Probably most of the words you don't know yet but you tell me if you know any of these. This one? This one? (E. quickly points to words in first column, then second column, etc.)

VISUAL METHOD

Materials: Stimulus Picture, Sentence Card, Word Picture Cards, Student Answer Booklets.

Directions:

A. TEACHING (8 minutes)

1. (Show stimulus picture.) Today you're going to learn to read some (more) words. Now the first thing I want you to do is to look at this picture and tell me what you see. (pause for r.) I'm going to make a story about what you see in the picture.

2. (Point to figure in picture.) Who (what) do you see here? (pause for r.) Good. (Point to first sentence on Sentence Card.) Here's a sentence that says, "I see a (E. says word)." I see a (Pause for r.). Good. This word that is underlined is a word that you're going to learn. Now look at it carefully. The word you're going to learn is (pause for r.) (E. says word.)

Repeat 2 for the remaining four words.)

3. (Review sentences on Sentence Card in mixed order) Now I'm going to say the first part of the sentence. You say the word when we come to it if you can. I see a (pause for r.). (Point to underlined word. If children don't know word E. says it). I see a (pause for r. to second word, E. pointing to it, saying it if necessary.)

(Repeat all 5 sentences a second time in mixed order.) Let's try that again. I see a (pause for r.)

4. (E. writes words on board or paper.) Look here at these words. Which one is the longest? (E. points with finger) Good. Does anyone know what word this is? Yes (or No. It's (E. says word). Remember, the longest word is (pause for r.). Which is the shortest word? (E. points with finger) Good. Does anyone know what this word is called? Yes (or No. It's (E. says word). Remember, the shortest word is (Pause for r.).

5. Now, let's make a house for each word. You watch and I'll show you how. I'm going to draw a line around the shape of the word like this. (Ex. does so) to make a house. This part is the house. If a letter goes up here, it goes to the roof or top of the house. When a letter goes down here, we'll call it the basement.* (Before doing next word: Do you think this next word will have a house like the first word? (pause) Let's see. (Cont. with other words in same way, describing where the letters "are going" as you make the house.)

*[Do you remember what the word with this shape is? (Pause for r.) Good/No it's _____.]

6. (Review all words.) The longest word is (pause for r.). The shortest word is (pause for r.) (For remaining words) This word is (pause for r.).

B. FIRST TEST TRIAL

1. (Give a response booklet to each child.) (E. writes child's name on booklet.) I am going to ask you to put a circle around some words. When you finish, cover up your answer like this.

All right, now look here at the first long box on top with five words in it. (E. demonstrates.) Put your finger on it. Now put a circle around the word. (E. reads first word from Test Trial One list.) (Repeat word.) Now don't answer out loud. Is it this? or This? or This? (E. points to words) Now put a circle around (E. says word.) All right, next put your finger on the second long box. Put a circle around (E. reads second word on list.) (Repeat word.) (Repeat directions where necessary.)

(Continue for remaining three words.)

2. (E. Collects test booklets.)

C. TEACHING (3 minutes)

(Use Word-Picture Cards, first show word side only, then show picture.) (Here are the words you just learned. This one is (pause for r.). (Reveal picture.) Yes or No. This word is (E. says word.)

(Repeat all words a second time.)

(Show the word-only side of the Word-Picture Cards.) Who thinks they can find this word in the story? It has a line under it. (E. holds word near Sentence Card. If child has difficulty, give him a choice: Is it this word or this word?). Good. What is the word? (Pause for r.). Yes (or) No. It's (E. says word.) (If there are six children, repeat one word.)

(Repeat for the remaining four words cards.)

D. SECOND TEST TRIAL

1. (E. Returns response booklets to children) Turn the first page back like this. (Demonstrate) Remember to cover your answer as soon as you circle the word I say. Now put your finger on the first long box. Put a circle around the word (E. names word; use Trial Two list). Next put your finger on the second long box. Put a circle around (second word on list).

(Continue for remaining three words).

2. Now turn your booklet over and look up here.

E. REVIEW I

1. (Use Picture-Word Cards) (Show word side) This word is (Flip to picture. E. names word) (Do once for each of five cards.)

F. THIRD TEST TRIAL

1. Turn your booklet over and turn the top page back like this (Demonstrate) Remember to cover your answer as soon as you circle the word I say. Now put your finger on the first long box. Put a circle around the word (E. names word, use Trial Three List). Next put your finger on the second long box. Put a circle around (second word on list).

(Continue for remaining three words).

G. REVIEW II

1. (Same as Review I.)

H. FOURTH TEST TRIAL

1. (Same as Third Test Trial.)

I. FIFTH TEST TRIAL

(Have each child come to you individually. Show him the word side of the Picture Word Cards.) What word is this? (Repeat for remaining four words. Record response on fifth page of answer booklet.) (mix up cards)

TEST TRIAL WORDS

VISUAL FORM I

Trial 1	Trial 2	Trial 3	Trial 4
children	tree	girl	house
tree	girl	house	bus
girl	house	bus	children
house	bus	children	tree
bus	children	tree	girl

VISUAL FORM 2

Trial 1	Trial 2	Trial 3	Trial 4
rabbit	farm	duck	wagon
farm	duck	wagon	boy
duck	wagon	boy	rabbit
wagon	boy	rabbit	farm
boy	rabbit	farm	duck

AUDITORY METHOD

Materials: 6 Letter-Sound Cards, 5 Word Cards, Word Family Card.

Directions:

A. TEACHING 9 minutes.

1. (Use Letter-Sound Cards. Present in numbered order) Today you're going to learn to read some (more) words. First we'll learn some sounds and then put them together into words.

2. (Show card 1, point to letter) This letter has a special sound. You can hear it at the beginning of (E. names first and second picture.) All right, this letter (E. traces letter) has the sound (E. gives sound.) Can you say it? (Pause for r.) Good. Can you hear (E. says sound) at the beginning of (E. says key word)? Now listen carefully. Does (E. says first word on back of card) begin with the

same sound as (E. says key word)? (E. repeats words together.) Good/Listen again carefully (E. repeats words exaggerating initial sound.) Does (second word on back of card) begin with the same sound as (E. says key word, repeats words together.) Does (third word on back of card) begin with the same sound as (E. says key word, repeats words together). Now remember, this letter has the sound (pause for r.) (E. gives sound) as in (E. says key word.)

(Repeat 2 for the remaining 5 Letter-Sound Cards.)

3. (Review six Letter-Sound Cards) Do you remember what sound this letter has? (Pause for r.). Yes (or) No. It has the sound (E. gives sound) as in (E. Names picture 1) and (E. Name picture 2). (Continue for remaining five cards.)

4. (Use word cards, pointing to the letters on top for blending) This sound is (pause for r.) (E. Names sound). This sound is (pause for r.). Now can you put these sounds together to make a word (E. gives sound) (gives sound) (pause for r.). (E. gives sound (gives sound) (E. Names word) like in (E. reads sentence on back of card.) "Let's see if you can put the sounds together now." (Pause for r., E. names sounds if necessary, and says, Put it together into a word) (pause for r.) (E. points to beginning and ending of word and whole word while saying sound).

(Repeat 4 for the remaining four blended word cards)

B. FIRST TEST TRIAL

1. (Give a response booklet to each child.) (E. writes child's name on booklet.) I am going to ask you to put a circle around the word. (E. reads first word from Test Trial One list.) (Repeat word.) (Watch to see that children mark in the right place. Give further directions if necessary.) All right, next put your finger on the second long box. Put a circle around (E. reads second word on list.) (Repeat word.)

(Continue for remaining three words.)

2. (E. Collects test booklets.)

C. TEACHING 3 minutes

1. (Use Letter-Sound Cards. Present in numbered order.) Do you remember what sound this is? (pause for r.) You can hear it at the beginning of (E. names first picture) and (E. names second picture). All right this (E. traces letter form) has the sound (E. gives sound.)

(Repeat 1 for the remaining 5 Letter-Sound Cards.)

2. (Review six Letter-Sound Cards.) Do you remember what sound this letter has? (Pause for r.). Yes/No. It has the sound (E. gives sound) as in (E. names picture 1) and (E. names picture 2). (Continue for remaining five cards.)

3. (Use Word Family Card.) (E. says all words emphasizing endings.) Now listen while I say these words. (E. says all words emphasizing endings.) See they all end the same. They all end with (E. names ending, pointing first to ending in box, then down to ending in each word.) Listen, can you hear it? (E. repeats 4 words and points as before). Do they begin with the same sounds? (pause for r.) No, they begin with different sounds. What does this one begin with? This one? This one? But remember, they all end with (pause for r.) (E. names ending and points to ending in box on chart.)

SECOND TEST TRIAL

1. (E. Returns response booklet to children.) (Turn the first page back like this. (Demonstrate) Remember to cover your answer as soon as you circle the word I say. Now put your finger on the first long box. Put a circle around (second word on list).

(Continue for remaining three words)

2. Now turn your booklet over and look up here.

E. REVIEW I

1. (Use Word Cards.) This word is (E. sounds out word, pointing to the beginning and ending and whole word.) (Do once for each of the five cards.)

F. THIRD TEST TRIAL

1. Turn your booklet over and turn the top page back like this (Demonstrate) Remember to cover your answer as soon as you circle the word I say. Now put your finger on the first long box. Put a circle around the word (E. names word, use Trial Two List). Next put your finger on the second long box. Put a circle around (second word on list).

(Continue for remaining three words.)

G. REVIEW II

1. (Same as Review I)

H. FOURTH TEST TRIAL

1. (Same as Third Test Trial.)

I. FIFTH TEST TRIAL

(Have each child comes to you individually. Show him the word side of the Picture Word Cards.) What word is this? (Repeat for remaining four words. Record response on fifth page of answer booklet.) (Mix up cards)

AUDITORY FORM 1

Trial 1	Trial 2	Trial 3	Trial 4
at	cat	hat	bat
cat	hat	bat	sat
hat	bat	sat	at
bat	sat	at	cat
sat	at	cat	hat

AUDITORY FORM 2

Trial 1	Trial 2	Trial 3	Trial 4
an	pan	fan	man
pan	fan	man	ran
fan	man	ran	an
man	ran	an	pan
ran	an	pan	fan

RETENTION TEST

Materials:

Answer booklet, test cards of Auditory and Visual Task words.

Directions:

(Work with each child individually; administer twenty-four hours after learning task.)

A. Production Method

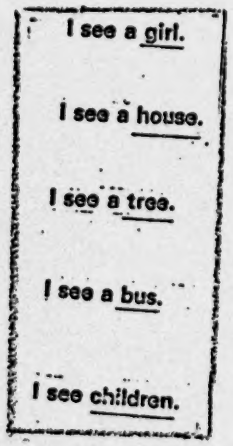
(Show the word side of the card) What word is this? (E. does not tell child whether or not his answer is correct. Record response on fifth page of answer booklet. Repeat for remaining four words.)

B. Recognition Method

(E. places all five cards word-only side up on the table.) Which word is (E. says one word)? (E. does not tell child whether or not his answer is correct. Record response on fifth page of answer booklet. Repeat for remaining four words.)

VISUAL ---------- EXAMPLES FROM MATERIALS ---------- AUDITORY

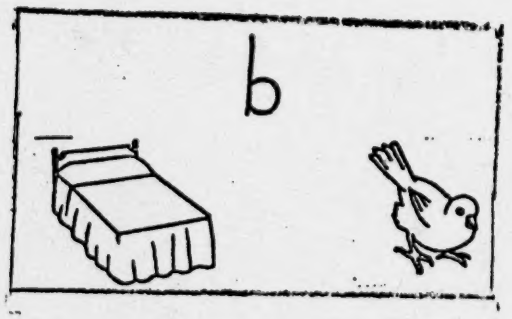
Stimulus Picture

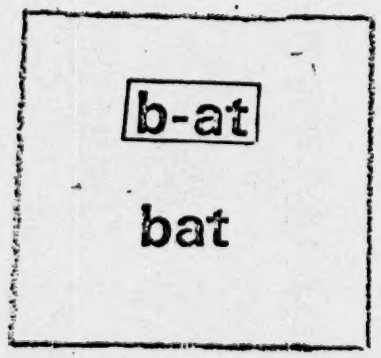


Sentence Card

Letter-Sound Card

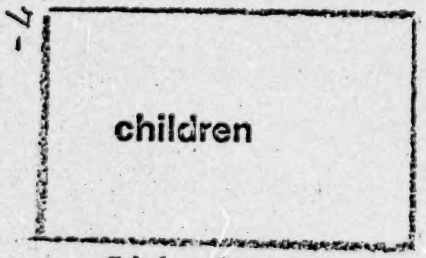
Comparison words:
bed - boy
bed - train
bed - bike





Word Card

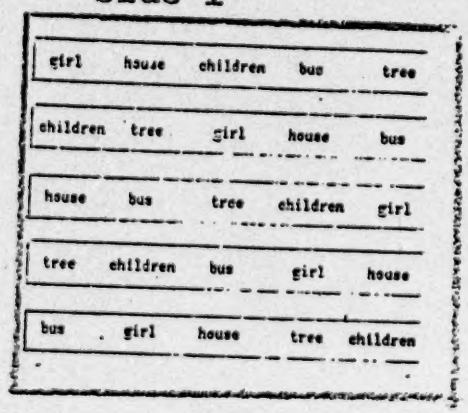
Word Family Card



Side 1

Word Picture Card

Side 2

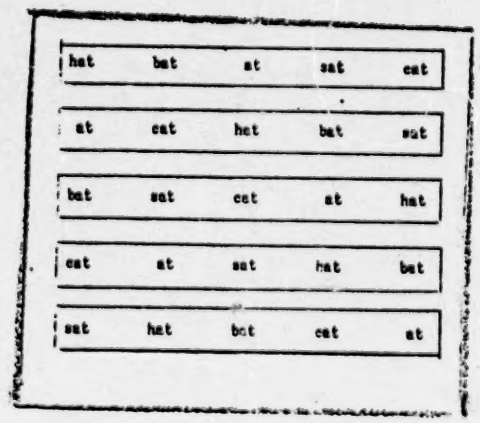

girl house children bus tree
children tree girl house bus
house bus tree children girl
tree children bus girl house
bus girl house tree children

Student Answer Booklet*

Summary sheet is shown on page 6.

*Pages 2-5 of the Booklet have these formats.

Student Answer Booklet*



APPENDIX B

MILLS LEARNING METHODS TEST

Mills Learning Methods Test

The Mills Test was designed so that four learning methods could be used and compared. For the purposes of the present study only two methods were used. Materials, procedures, and directions used are as follows:

Test Materials

One set of picture word cards, the primer level, which were devised by Mills were used. Any words also presented on the Tasks were excluded from this sample. The cards are 4" x 6", and similar to the Dolch Picture Wards.

Procedures

Mills recommends that for beginning readers, 20 words be randomly selected and half be assigned to each of the auditory and visual methods. Teaching for each method took 15 minutes; following the teaching the child was asked to name the ten cards. The number correctly named was his training score.

Twenty-four hours later, retention is checked by having the child name the ten words.

Directions:

SPECIFIC TEACHING PROCEDURES[1]

"In order to obtain specific techniques or steps that would be typical of a particular method of teaching word recognition, activities were selected for each method on the basis of the frequency of mention in the literature available regarding the respective methods. It is necessary to use these standard teaching procedures in order to make a valid comparison between the various methods.

VISUAL METHOD

Using ten 'unkown' words from the controlled list provided, teach the child recognition of these words stressing exclusively the visual appearance and other visual clues of the words through the use of the following procedures for 15 minutes:

---

[1]Mills, loc.cit.

1. Present the ten picture-word cards to the child with picture-side up. Ask him to look at the picture, then at the words, and to say what it is. Do this with all ten of the cards.

2. Present the cards a second time with picture-side up. This time show the picture and the words and have the child make a sentence using the word.

3. Play a game of solitaire with the child. Do this by placing all ten of the cards with word-side up before the child. As he says each one of them, ask him to turn the card over to check his accuracy with the picture.

4. Ask the child to arrange the words in a meaningful sequence to form a sensible phrase or sentence. In some instances he will be able to use three or four of the words in a single sentence. He may add any other necessary words to make a complete sentence.

5. Discuss the relative length of various words. Print the words on the blackboard or on paper, and then ask the child the following questions: Which is the longest word? Which is the shortest? Have him match words which are of the same length. Ask him to arrange the words in groups by stacking all the three letter word cards in one stack, the four in another stack, etc.

6. Draw on the board (or paper) a figure, the outlines of which follow the shape or gestalt of each one of the words. For example, a box or figure depicting the word form would be tall at the beginning following the height of the letter f and then would take a long, low rectangular form for the remainder of the word. After all ten shapes have been drawn, ask the child to match these various shapes with the words for which they stand.

7. Repeat activity "1" as described above.

8. If after activity "7", fifteen minutes have not elapsed since the lesson began, continue repeating the steps in order until exactly fifteen minutes have been spent on the lesson.

9. At the end of the fifteen minutes, administer the immediate recall test. Do this by exposing the word-side of the card to the child. If he responds correctly within the five seconds of exposure of the word, record the right response by placing + on the test record form by that word in the "Immediate" column in the space provided. Do this for all ten words. Record the total number of correct responses.

PHONIC OR AUDITORY METHOD

Using ten new "unkown" words from the controlled list provided, teach the child recognition of these words, stressing exclusively the sound qualities of the words through the use of the following teaching procedures:

1. Print the words on the blackboard (or paper), then say the word slowly with the child repeating the word after you. Ask the child to make a sentence with the word to insure that he understands the meaning. Then point to the opening letter of the first word. Say, "This is the letter___. It makes sound such as ___. The next letter is the letter____; it makes the sound of ___, etc." Have the child repeat the separate sounds after you as you do them with each word. Then have the child try the entire word sound by sound, then blending the different sounds into one whole until it is a unified, complete word. Do as much of this as possible in the first four or five minutes of the lesson.

2. Ask the child to think of other words beginning with the same sound as the first sound of ___ or words ending with the same sounds, or words having the same middle sound as the one being studied. Have the child arrange the cards (never present the picture-side of the cards during this phonic lesson) in different groups as he identifies and classifies the ten words as to common beginnings or endings.

3. Help the child to think of words that rhyme with the word you are trying to teach him.

4. Have the child identify familiar sounds or little-words-in-big-words of those being studied.

5. Repeat all ten words for the child with exaggerated sound stresses, with short pauses between the sound elements. Ask the child to listen closely to the sounds as you point to the particular part of the word being sounded. Then ask the child to say the word.

6. Ask the child to say the words now without your help. Have him listen closely to the sounds as you point to the letters as he slowly pronounces the words. (Do help the child with those words in which he still lacks recognition).

7. If fifteen minutes have not elpased since the lesson began, continue repeating the steps in the order listed above until exactly fifteen minutes have been spent on this lesson.

8. At the end of the fifteen minutes, administer the immediate recall test for these ten words."

APPENDIX C
READING AND READINESS INSTRUMENTS

READING AND READINESS MEASURES

In this section information is presented concerning the validity and reliability of the Gates-MacGinitie Reading Tests[1] and the Wide Range Achievement test.[2] The form used in the evaluation of December and May reading (Words Learned, Phonics, and WRAT) and the kindergarten teacher rating form are shown.

Reliability coefficients (split half) reported for Level I of the Wide Range Achievement Test (WRAT) ranged from .981 to .993. Reliability coefficients estimated by comparing Levels I and II ranged from .883 - .936. The degree of relationship between the WRAT and teachers' rating of achievement and between the WRAT and mid-term grades for a fifth grade sample was substantial (.78 and .88, respectively.)

Similarly, reliability coefficients for the Vocabulary and Comprehension sections of the Gates-MacGinite Reading Tests, Primary A, was reported as .91 and .94 respectively using a split-half method of estimation and some what lower, .86 and .83, respectively, using alternate forms. Validity coefficients estimating the relationship between the Gates-MacGinitie. (Forms D and E only) and the Lorge-Thorndike Verbal IQ ranged from .60 for fourth graders to .86 for seventh grade children.

---

[1]Gates, A. I., & MacGinitie, W. H., loc. cit.

[2]Jastak, J. F., Jastak, S. R., loc. cit.

ANSWER FORM--MAY READING

Score_____

Name_____ No. taught_____

School_____ Efficiency_____

Teacher_____ WRAT_____

Book_____ Gates V._____ C._____

Page_____ Nonsense Words_____

Regular List (=8)	Correct I	Correct D	Error ( ) NR

Sup. List (=4)	Correct I	Correct D	Error ( ) MR

Phonics List	Correct I	Correct D	Error NR
lat			
wub			
con			
jed			
dit			
po			
te			
fi			
ra			
mu			
bane			
sebe			
hime			
node			
gute			

WRAT____ LEVEL I____ Correct = NR + X Error - record Ceiling = 12 words

cat see red to big work book eat was him how 36

then open letter jar deep even spell awake block size 46

weather open lip finger tray felt stalk cliff lame struck 56

Write name; name 2 letters A R Z H I Q S E B O

_____ A B O S E R T H P I U Z Q 25

TEACHER RATING SCALE

Student	Birthdate	Sex	Speech Difficulty
Kindergarten teacher	Parent Occupation		Pre-test score

How would you rate the child on the following characteristics in comparison with others in your class: Encircle the number which most appropriately describes the child on each characteristic. Feel free to make any additional comments on the bottom of the page.

1. Verbal Fluency
   Expression of Ideas

   Superior 1 2 Average 3 4 Poor 5 6

2. Non-Verbal
   Problem Solving

   Superior 1 2 Average 3 4 Poor 5 6

3. Attention Span
   Ability to Complete a Task

   Highly Attentive 1 2 Average 3 4 Easily Distractable 5 6

4. Speech, Articulation
   Pronunciation of Words

   Clear, Accurate 1 2 Average 3 4 Difficult to Understand 5 6

5. Listening Comprehension
   Following Directions

   Superior 1 2 Average 3 4 Poor 5 6

6. Visual Discrimination

   Superior L 2 Average 3 4 Poor 5 6

7. Motor Coordingation

   Well Coordinated 1 2 Average 3 4 Awkward Immature 5 6

8. Reading Readiness

   Very Ready 1 Average Readiness 2 Needs Work 3 Considerable Work Needed 4

9. By which method do you predict the child will most easily learn to read?

   Auditory 1 Visual 2 Either 3 Neither 4

COMMENTS:

APPENDIX D
NUMBER OF SUBJECTS SELECTED AND TESTED

Table 29 shows the number of children, divided according to sex, area, time of testing on the Tasks, and test comparison group, who participated in the study. The number of children selected within each group, the number with four learning task measures and four retention scores, and the number with complete Grade 1 reading data is summarized in the Table.

TABLE 29

NUMBER OF SUBJECTS IN EACH SUBGROUP (1) SELECTED FOR STUDY, (2) WITH COMPLETED LEARNING TASK DATA, AND (3) WITH COMPLETE FIRST GRADE READING DATA.

Test Comparison Group	Urban May		Sub. May		Urban Sept.		Sub. Sept.		Total	
	M	F	M	F	M	F	M	F	M	F
Form 1 - Form 2										
(1) Selected	12	12	12	12	9	9	12	12	45	45
(2) Complete Learning Task Data	12	7	12	11	18	7	11	11	43	36
(3) Complete Reading Data	10	4	10	10	6	7	10	10	36	31
Group-Individual										
(1) Selected	12	12	12	12	9	9	12	12	45	45
(2) Complete Learning Task Data	10	12	10	9	8	9	12	12	40	42
(3) Complete Reading Data	8	11	10	8	7	8	12	12	37	39
Task-Mills										
(1) Selected	-	-	-	-	9	9	12	12	21	21
(2) Complete Learning Task Data	-	-	-	-	8	7	12	11	20	18
(3) Complete Reading Data	-	-	-	-	8	7	12	11	20	18
Total Complete	18	15	20	18	21	22	34	33	93	88

188 93

APPENDIX E
IDENTIFICATION OF DISABLED READERS

Graphic descriptions of the relationship between scores on the tasks, the Metropolitan and Words Learned-December, respectively, and May reading scores (WRAT) are shown in Figures 6-8.

FIGURE 6

DISTRIBUTION OF SCORES ON THE TASK
IN RELATION TO SCORES ON THE WRAT

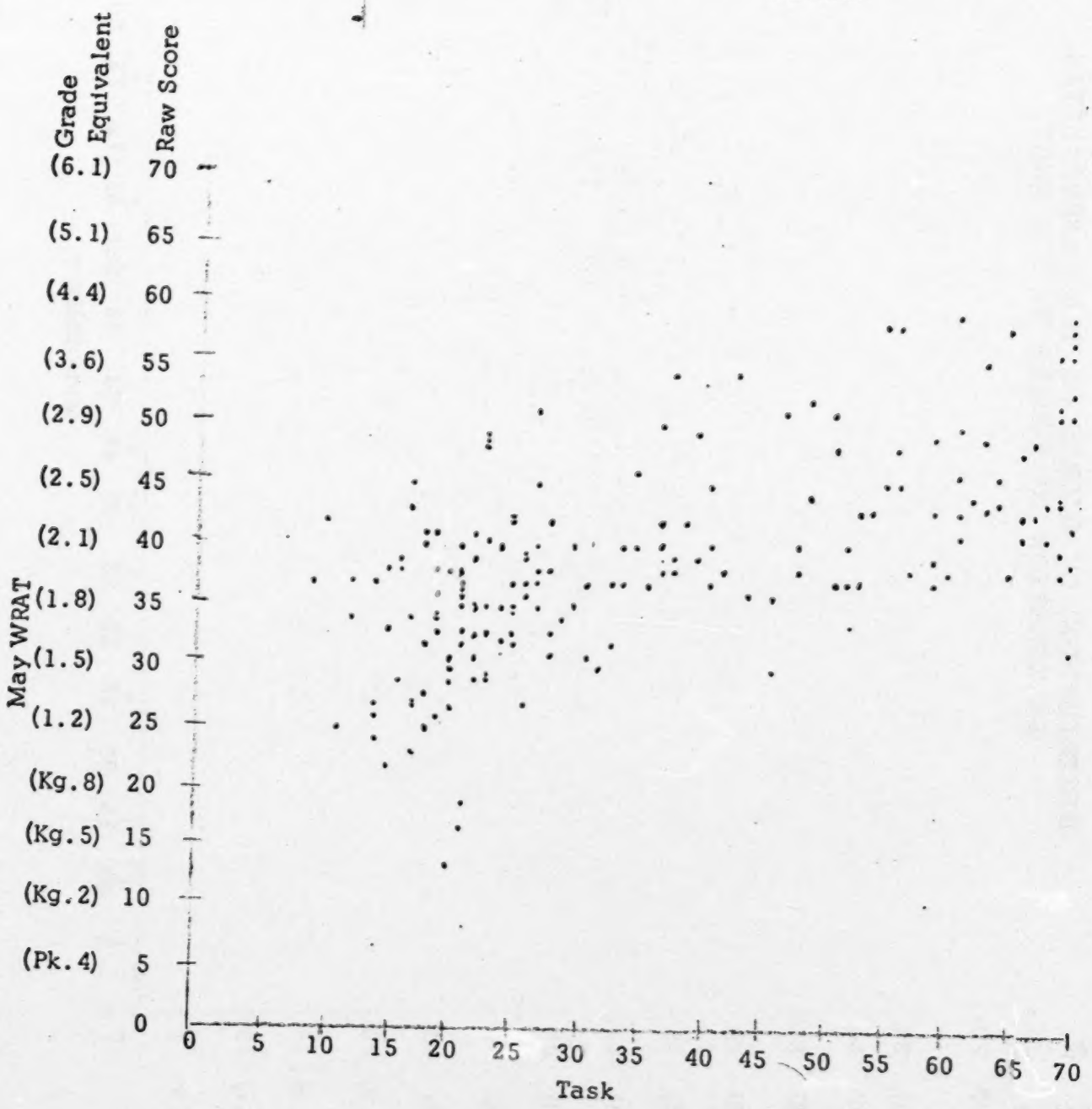


FIGURE 7

DISTRIBUTION OF SCORES ON THE METROPOLITAN
IN RELATION TO SCORES ON THE WRAT

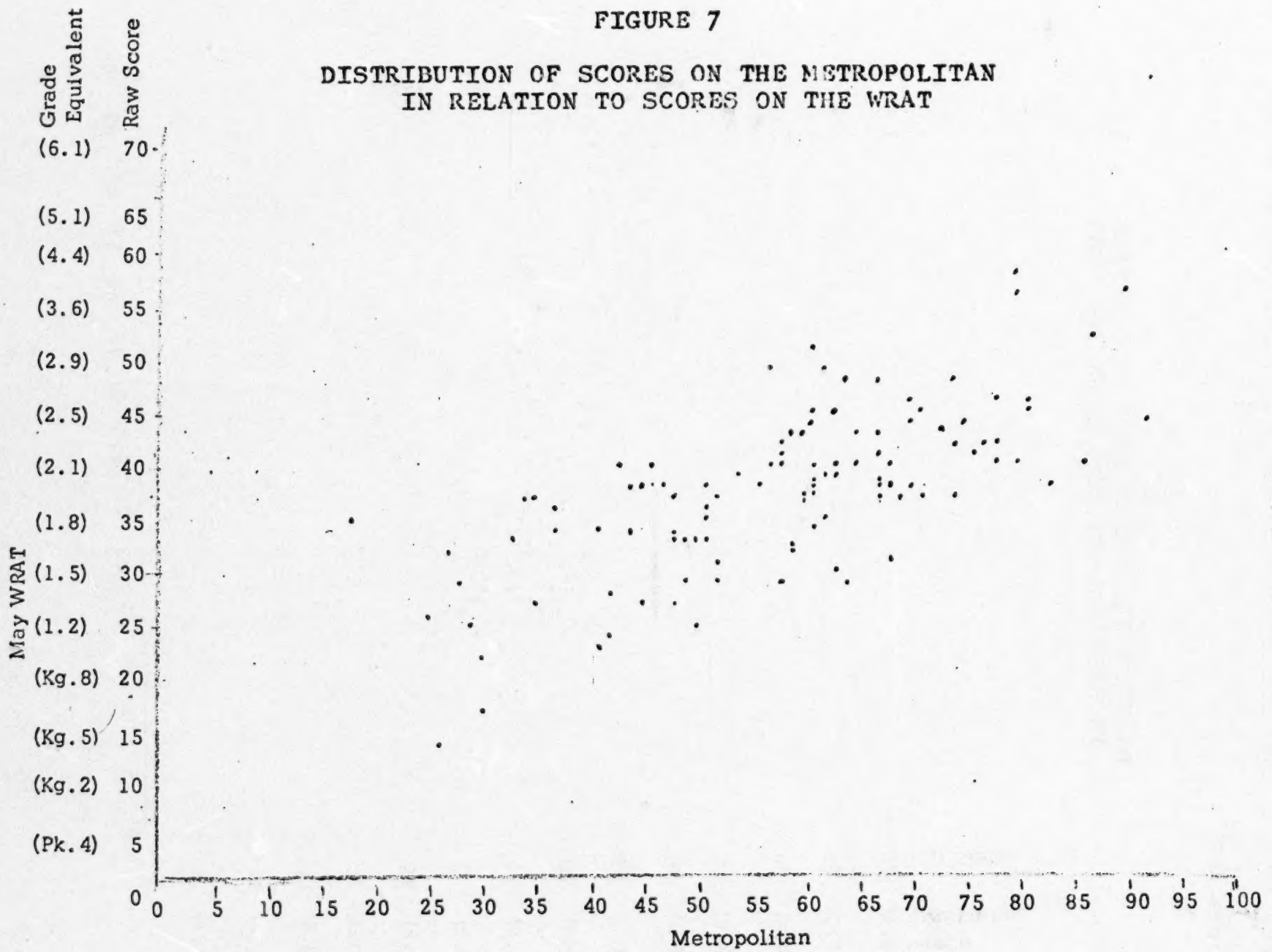


FIGURE 8

DISTRIBUTION OF SCORES ON THE WORDS LEARNED-DECEMBER IN RELATION TO SCORES ON THE WRAT

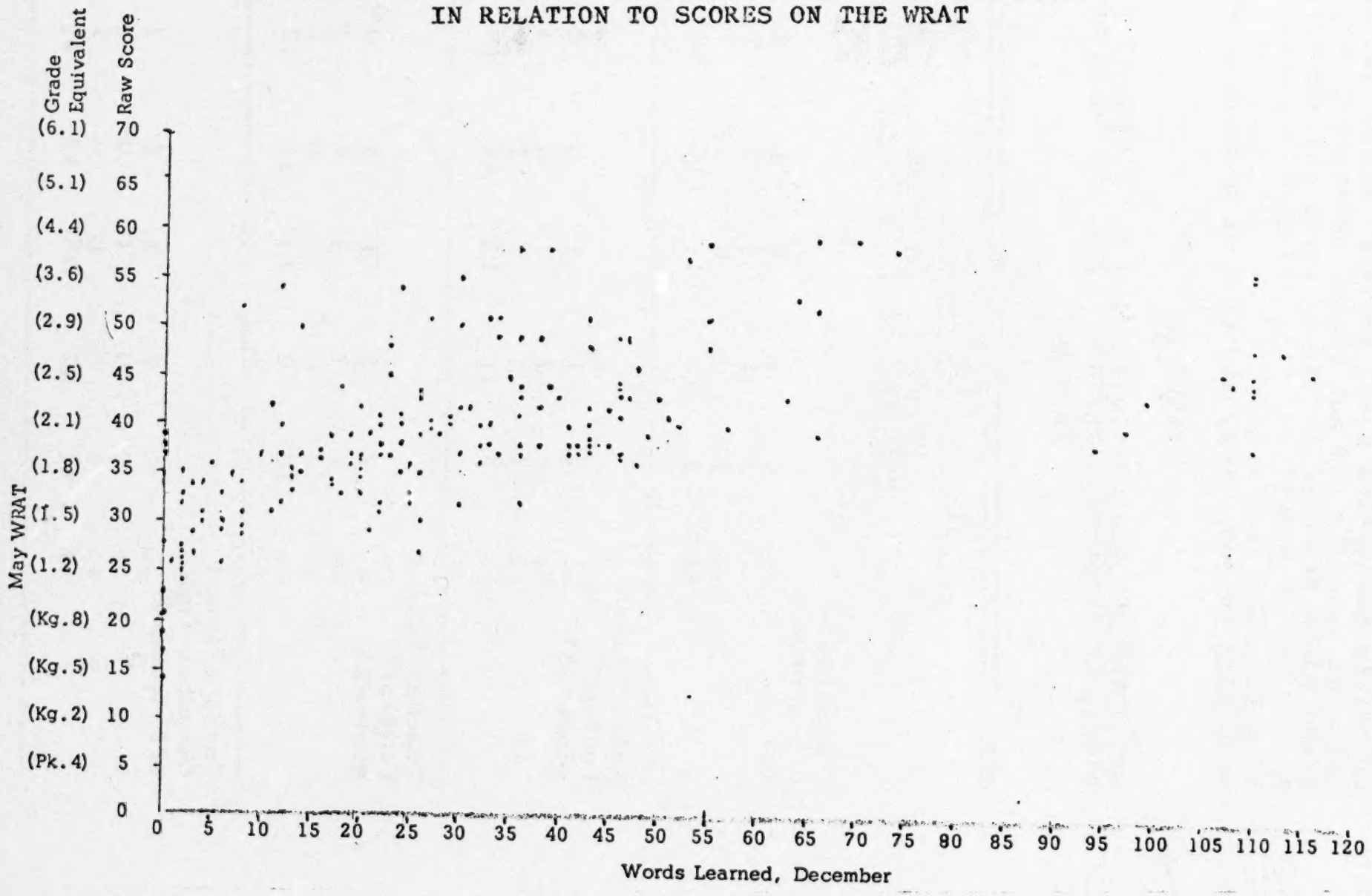


Table 30 shows the number of disabled readers correctly identified (--), children identified as needing help who actually did not (false negatives, -+), children found to be disabled in May of grade one who were overlooked (false positives, +-), and children correctly identified as good readers (++). Subjects are grouped according to sex, area, and time of testing on the tasks.

TABLE 30

SUBGROUPS OF SUBJECTS IDENTIFIED USING CUT-OFF POINTS ON THE TASKS, METROPOLITAN, TEACHER RATING AND WORDS LEARNED-DECEMBER. (N = 108)

	Group[1]	May Sample				September Sample		
		Urban		Suburban		Urban		
		Male	Female	Male	Female	Male	Female	Total
Tasks (cut-off score=20)	--	6	2	1	1	2	1	13
	-+	4	1	5	2	1	4	17
	+-	1	0	0	0	2	0	3
	++	7	11	14	15	14	14	75
Metropolitan (cut-off score=43)	--	5	1	1	0	4	1	12
	-+	2	1	4	0	3	7	17
	+-	2	1	0	1	0	0	4
	++	9	11	15	17	12	11	75
Teacher Rating (cut-off score=2)	--	7	2	1	1	0	0	11
	-+	1	3	5	1	3	1	14
	+-	0	0	0	0	4	1	5
	++	10	9	14	16	12	17	78
Words Learned December (cut-off score=3)	--	6	2	1	0	3	1	13
	-+	0	0	1	0	0	2	3
	+-	1	0	0	0	1	0	2
	++	11	12	18	18	15	16	90

[1]Interpretation of signs is given in Table 25, page 54.

REFERENCES

Bagford, J. Reading readiness scores and success in reading. Reading Teacher, 1968, 21, 324-28.

Barrett, T. C. The relationship between measures of pre-reading visual discrimination and first grade reading achievement: a review of the literature. Reading Research Quarterly, 1965, 1, 51-76.

Birch, H. G. & Belmont, L. Auditory-visual integration, intelligence, and reading ability in school children. Perceptual and Motor Skills, 1965, 20, 295-305.

Bruce, D. J. The analysis of word sounds by young children. British Journal of Educational Psychology, 1964, 34, 158-70.

Budoff, M. & Quinland, D. Reading progress as related to efficiency of visual and aural learning in the primary grades. Journal of Educational Psychology, 1964, 55, 247-52.

Calfee, Robert et al. How a Child Needs to Think to Learn to Read. Technical Report No. 131. Madison, Wisconsin: Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, 1970.

Davidson, H. P. An experimental study of bright, average, and dull children at the four-year mental level. Genetic Psychology Monographs, 1931, 9, 119-289.

Deputy, E. C. Predicting First-Grade Reading Achievement. Contributions to Education, No. 426. New York: Teachers College, Columbia University, 1930.

Dukes, W. F. & Bastian, J. Recall of abstract and concrete words equated for meaningfulness. Journal of Verbal Learning and Verbal Behavior, 1966, 5 (October).

Dykstra, R. Auditory discrimination ability and beginning reading achievement. Reading Research Quarterly, Spring 1966, pp. 5-34.

Elkonin, D. B. The psychology of mastering the elements of reading. In B. Simon (ed.), Educational Psychology in the U.S.S.R. London: Routledge & Kegan Paul, 1963.

Gates, A. I. A study of the role of visual perception, intelligence, and certain associative processes in reading and spelling. Journal of Educational Psychology, 1926, 17, 433-45.

Gates, A. I. & Boehm, E. A study of initial stages in reading by pre-school children. Teachers College Record, 1923, 24, 469-88.

Gates, A. I. et al. Methods of Determining Reading Readiness. New York: Teachers College, Columbia University, 1939.

Giebink, J. W. & Goodsell, L. L. Reading ability and associative learning for children with a visuomotor deficit. American Educational Research Journal, 1968, 5, 412-20.

Hartley, R. N. Effects of list types and cues on the learning of word lists. Reading Research Quarterly, 1970, 6(1), 97-121.

Hill, M. B. A study of the process of word discrimination in individuals beginning to read. Journal of Educational Research, 1936, 29, 487-500.

Jeffrey, W. E. & Samuels, S. J. Effect of method of reading training on initial learning and transfer. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 354-58.

Linder, R. & Fillmer, H. T. Auditory and visual performance of slow readers. Reading Teacher, 1970, 24, 17-22.

MacGinitie, W. H. Evaluating readiness for learning to read: a critical review and evaluation of research. Reading Research Quarterly, 1969, 4, 369-410.

McNeill, J. D. & Stone, J. Note on teaching children to hear separate sounds in spoken words. Journal of Educational Psychology, 1965, 56, 13-15.

Muehl, S. & Kremenak, S. Ability to match information within and between auditory and visual sense modalities and subsequent reading achievement. Journal of Educational Psychology, 1966, 57, 230-39.

Murphy, H. A. Growth in perception of word elements in three types of beginning reading instruction. Reading Teacher, 1965, 19, 585-89, 600.

Roslow, S. Reading readiness and reading achievement in the first grade. Journal of Experimental Education, 1940, 9, 154-59.

Samuels, S. J. Attentional process in reading: the effect of pictures on the acquisition of reading responses. Journal of Educational Psychology, 1967, 58, 337-52.

Silver, A., Hagin, R., & Hirsh, M. Reading disability: teaching through elimination of deficit perceptual areas. American Journal of Orthopsychiatry, 1967, 37, 744-52.

Venezky, R. L. English orthography: its graphical structure and its relation to sound. Reading Research Quarterly, 1967, 2, 75-105.

Wiley, W. E. Difficult words and the beginner. Journal of Educational Research, 1928, 17, 278-89.